Essays in Applied Microeconomics: Job Mobility and Social Networks

Jan Sebastian Nimczik

Inauguraldissertation
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Abteilungssprecher:
Prof. Dr. Jochen Streb

Gutachter:
Prof. Dr. Andrea Weber
Prof. Antonio Ciccone, PhD

Tag der mündlichen Prüfung:

12. Oktober 2017

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig angefertigt und die benutzten Hilfsmittel vollständig und deutlich angegeben habe.

Mannheim, 15.7.2017

Jan Sebastian Nimczik

Curriculum Vitae

2012 – 2017	University of Mannheim, Department of Economics Ph.D. student in Economics
Fall 2015	University of California, Berkeley, Department of Economics Visiting Student Researcher
2007 – 2012	Albert-Ludwigs-University Freiburg Diplom-Volkswirt
2010 – 2011	University of Wisconsin-Madison, Department of Economics Graduate Exchange Program
2004 – 2010	Musikhochschule Freiburg and Sibelius Academy Helsinki Diplom-Musiker and Diplom-Musiklehrer

Für Caro & Clara.

Contents

Acknowledgments

First and foremost, I would like to thank my advisors, Andrea Weber and Antonio Ciccone, for their continuous support, invaluable feedback, and for always encouraging me to explore new ideas. I am glad about the opportunity to learn and benefit from their guidance and from interactions that shaped the way I am thinking about economics and research. I am particularly thankful to Andrea for providing access and introduction into the Austrian database that is fundamental for large parts of this thesis.

My sincere thanks go to my co-authors Albrecht Bohne (Chapter 2), Albrecht Glitz, Virginia Minni and Andrea Weber (Chapter 3), and Bernd Fitzenberger (Chapter 4) for the development and exchange of ideas, for many helpful discussions, and for a constructive working atmosphere. My research benefited from numerous comments and suggestions, in particular by Andreas Dzemski, Sebastian Findeisen, Jasper Haller, Andreas Peichl, Anna Raute, and Sebastian Siegloch. I would like to thank David Card for the opportunity to visit the University of California, Berkeley and his insightful comments and discussions. Furthermore, I had the opportunity to present the research leading up to this thesis at various seminars, workshops and conferences and I am grateful for the numerous comments and suggestions that participants at these events generously offered.

Very special thanks go to my friend, colleague, and co-author Albrecht Bohne for in- and extensive discussions during coffee breaks, for great company on our trips to Barcelona, Madrid and Ecuador, for our joint work efforts, and for being a great friend.

The main topic of this thesis is not least influenced by my personal experience of switching professions. I would like to thank my parents, Ortwin and Heike Nimczik, for their continuous support throughout these changes and all times.

Finally, I am very grateful to my wife Julia-Caroline Walther for her love, her never-ending support, and for the wonderful life with our daughter Clara Josephine.

Introduction

In many policy debates, flexibility of labor markets is postulated as a solution to high unemployment and low growth. An important task for academic research therefore is to analyze and understand the drivers and consequences, impediments and boundaries of job mobility. In this thesis, I examine job mobility from various angles with a particular focus on the role of network structures in determining worker transitions. While individual utility maximization is the predominant paradigm in most economic models, many decisions of economic agents are influenced by connections to other individuals, by social norms, and by larger groups. My dissertation aims to help understanding the impact of social interactions between firms and workers on economic outcomes and decisions such as job mobility.

In the first chapter, I explicitly analyze and model the formation of economic connections between firms as a stochastic process. I exploit observed worker transitions in order to identify the boundaries of job mobility and, hence, the boundaries of labor markets. To this aim, I propose a novel network-based method that allows to determine separate and largely self-contained labor markets.

In the second chapter, I examine how firm-level practices in taxation influence the behavior of individuals. In this context, worker flows between firms serve as a mechanism of information transmission and reduce information frictions. Notably, workers learn about incentives inherent in the tax system by switching to firms with a high knowledge of the system.

Informational uncertainties are also at the heart of the third chapter. In the context of hiring relations between firms, I show how learning processes through repeated interaction between firms can help to mitigate uncertainty about worker quality. This mechanism can lead to increased labor market efficiency by improving the quality of matches between workers and firms.

In the fourth chapter, I contribute to the range of econometric methods that deal with dependency structures in the data. The prevalence of inter-dependencies between observation invalidates standard independence assumptions and requires bootstrap methods that explicitly allow for correlation between error terms. I adjust the scheme for drawing bootstrap resamples in order to accommodate the specific dependence structure in the data.

Although the context of these studies is diverse, they share important common ground in their methodology. First, the empirical analysis in all chapters is based on detailed large-scale matched employer-employee data. Second, all studies extensively rely on the informational content embedded in worker flows. In recent years, labor economists increasingly explore the information embedded in worker flows between firms. The main idea is that job transitions between different employers reveal unobserved preferences that cause the job switch. At the same time, a job transition represents an expert assessment of the employer regarding the match of the worker to the skills required at the specific target firm. In the previous literature, worker mobility has predominantly been used to study unobserved worker and firm quality. In this thesis, I extend the range of applications that exploit the informational content of job mobility. Furthermore, I extend the methodology to analyze the data by incorporating methods from network analysis and machine learning. In the following, I provide a short introduction into the four different chapters of my thesis.

Labor Markets

In the first chapter, I develop a novel method that allows to determine the size and scope of labor markets within countries. A specific definition of labor markets within a country is a prerequisite for many applications in labor economics and related fields. Important studies in the current literature exploit variation across separate labor markets in order to answer key economic questions such as to identify the effect of global trade shocks, immigration, or technological change. For simplicity, most studies approximate labor markets by geographical regions and administrative boundaries. These geographically separated *local labor markets*, however, are subject to a number of important drawbacks. First, empirical researchers have little guidance which is the geographical entity that forms the relevant reference group for firms and workers. Second, administrative boundaries are relatively stable over time while modern technology such as online job search has vastly increased the potential size of the relevant labor market for job seekers. Moreover, workers are heterogeneous in their preferences for mobility and face large differences in the local availability of jobs in their profession.

In order to address these issues, I use modeling techniques from the literature on community detection in complex networks to identify groups of firms that form *endogenous labor markets*. Unlike existing definitions, the novel concept does not rely on predefined geographical boundaries. Instead, labor markets are revealed by common patterns in observed job mobility flows across firms. In particular, firms are classified in the same market if they have similar probabilities of job flows to other firms. Based on large-scale matched employer-employee data from Austria, I show that the resulting labor markets are spatially clustered but deviate from administrative boundaries in various aspects. Notably, the approach reveals several separate markets within geographical areas as well as some labor markets that are scattered around the entire country. Moreover, the structure of labor markets changes over time as increasing mobility widens the geographical scope of labor markets.

The flexibility of this data-driven determination of labor markets allows for a range of new economic insights. I demonstrate that employment spillovers following a large local labor demand shock are not necessarily a local phenomenon but also affect distant firms from the same estimated labor market. Notably, the breakdown of the Austrian steel industry at the end of the 1980s caused spillovers on employment in distant non-steel related firms from the same endogenous labor market but left employment in other labor markets unaffected. Further, I show that worker reallocation in response to global trade shocks takes place across regions and industries but within endogenous labor markets. Summarizing the evidence, I find that endogenous labor markets capture the relevant reference group for estimating spillover effects and mobility responses to local labor supply and demand shocks.

Behavioral Responses to Taxation

The second chapter is based on joined work with Albrecht Bohne. In this chapter, I examine how individuals respond dynamically to incentives provided by the tax system. Most importantly, I analyze how and from whom workers learn about opportunities to avoid paying taxes as they gain experience in the formal tax system. Specifically, I document the important role of firms and job mobility in transmitting information about the tax system.

Most countries deploy a progressive schedule of personal income taxes where the tax rate jumps discretely upwards for income above certain thresholds. The standard theory of labor supply examines the trade-off between labor and leisure and predicts that some individuals adjust their income to values just below these kinks in the tax schedule. The empirical evidence for such *bunching* responses to income taxation, however, is limited. On the one hand, this can be due to adjustment frictions as wage contracts offered by firms are not flexible to be adjusted to personal needs. On the other hand, it can be due to information frictions as individuals might not be aware of the incentive structure and tax avoidance opportunities.

I analyze the information channel in a particularly well suited setting in Ecuador. Like many developing countries, Ecuador is in the transition from an informal to a more formal economy. The government provides a broad range of incentives and tax rebates to induce workers to switch into the formal tax system. In particular, the Ecuadorian tax law enables individuals to deduct personal expenses for housing, clothing, education, health and nutrition on a very large scale.

Based on unusually rich data that covers tax return data on the universe of tax payers in Ecuador and matched socio-demographic data from the civil registry and the firm registry, I examine how taxpayers increasingly exploit possibilities to avoid taxes with growing experience in the tax system. Following individuals over time, I show a strong increase in the probability to report earnings in the vicinity of the first kink in the Ecuadorian tax schedule. This bunching behavior is predominantly driven by the use of the generous deduction opportunities.

A closer look on the determinants of the increase in bunching reveals the importance of job mobility for the transmission of knowledge about tax avoidance opportunities. Workers learn about tax incentives and the possibility to avoid tax payments by switching to firms with a high share of bunchers. Consistent with a story of learning and memory, workers continue to bunch when switching to firms with lower levels of bunching. In contrast, individual workers entering a new firm have little influence on tax avoidance behavior of their incumbent co-workers. In conclusion, the firm is the relevant reference group that shapes individual responses to income taxation. Learning about the system is mainly driven by movements to specific firms that provide workers with information about tax avoidance.

Sourcing, Learning, and Matching in Labor Markets

In the third chapter, which is joint work with Albrecht Glitz, Virginia Minni, and Andrea Weber, I analyze hiring relations between firms as an important determinant of job mobility.

For firms, it is an important task to hire the right, i.e., highly productive, workers. This is emphasized by the huge effort most firms undergo when deciding whom to hire. In many dimensions, however, worker quality is unobserved and can be screened only imperfectly. Hiring decisions are therefore accompanied by high levels of uncertainty. In order to mitigate this problem, firms can rely on various types of networks that help to reduce uncertainty around worker quality. While the literature has started to explore the role of personal relations in reducing uncertainty through referrals (for instance by former co-workers or individuals from the same ethnic group), there has been much less focus on the role of networks among firms directly. In particular, firms gain experience from repeated interaction with each other and might be able to learn about the match quality of workers from particular firms.

I document a range of empirical facts that emphasize the importance of experience of hiring firms with specific source firms for hiring decisions. First, I show that older firms tend to poach their workers from a narrower set of source firms. As the age of a firm increases, its hiring gets more concentrated and new hires come from a smaller number of selected firms when conditioning on firm size and growth. This result is robust when controlling for other firm characteristics, industry classes and geographical factors. Second, I show that a firm's acquired experience in hiring from a specific source firm leads to higher starting wages and longer tenure of workers hired from that particular firm. Having gained experience from previous interactions with the same source reduces uncertainty and therefore results in better matches. With increasing tenure, however, the information advantage disappears and workers that are hired from sources with no previous experience catch up.

To substantiate these empirical findings, I build a search and matching model with heterogeneous workers, on-the-job search and match-quality as a pure experience good. In the model framework, a firm learns over time about the average quality of workers from different source firms. If a firm repeatedly proves to be a good source of highly productive matches, hiring firms are more willing to employ workers from this particular source. As a result, new workers hired from the most appropriate source firms are, on average, better matched to their firm than workers hired from other firms. Hence, they earn higher starting wages and have lower separation rates. However, as workers and firms learn about their idiosyncratic match-specific productivity, low quality matches are terminated and the wage and turnover advantage dissipates over time.

Clustered Standard Errors

The fourth chapter grew out of my Diploma Thesis at the Albert-Ludwigs University of Freiburg and is joint work with Bernd Fitzenberger. In this chapter, I propose a novel bootstrap method to deal with dependency structures in the error term of linear regression models when there is only a small number of clusters. It is well understood that few clusters cause problems in the convergence of standard variance estimators since asymptotic results are based on the assumption that the number of clusters goes to infinity. Bootstrap approaches are a common solution to this problem. It is however less known that few clusters also cause problems if the data set contains cluster-invariant binary variables. In this case, the probability that certain bootstrap resamples contain only zeros or ones is relatively high and leads to perfect multicollinearity. Standard software packages simply ignore this problem and report unreliable results. A modified bootstrap scheme that guarantees that each resample contains every cluster in the data set but attaches randomly generated weights to these clusters overcomes this issue. I provide extensive simulation evidence for the favorable performance of the weighted bootstrap method. Drawing weights from a uniform distribution and correcting the estimator for the variance of this distribution leads to a minimal but persistent asymptotic refinement on standard pairs cluster bootstrap methods.

Chapter 1

Job Mobility Networks and Endogenous Labor Markets

1.1 Introduction

Large parts of the current literature in labor economics rely on the concept of separate and largely self-contained local labor markets. Recent examples include studies that use variation between local labor markets to identify the impact of global trade shocks (Autor, Dorn, and Hanson, 2013) or immigration (Dustmann, Schönberg, and Stuhler, 2016) on wages and employment. Boundaries of local markets are also important to determine treatment and control groups in the evaluation of local policies and shocks (e.g., Lalive, Landais, and Zweimüller, 2015) and in the analysis of spillover effects (Crepon, Duflo, Gurgand, Rathelot, and Zamora, 2013; Ferracci, Jolivet, and van den Berg, 2014). Furthermore, the definition of local markets is crucial for the explanation of disparities in regional economic activity within countries, agglomeration economies, and the estimation of spatial equilibrium models (for an overview see Moretti, 2011).

From an empirical perspective, however, it is unclear how to determine the boundaries of local markets. More generally, it is unclear whether geographical borders of labor markets can be identified at all. The predominant approach in the literature uses predefined, geographically separated regional entities such as states, metropolitan statistical areas, or counties to approximate labor markets. In a more elaborate concept, commuting zones pool smaller areas that are connected through high commuter flows. These concepts however are subject to a number of important drawbacks. First, empirical researchers have little guidance on which specific geographical unit to consider.[1] Second, secular trends in the geographical mobility of

[1]For instance, Moretti (2011) provides a discussion of human capital spillovers on wages within local labor markets. He argues that differences in the evidence for spillovers on the state level (Acemoglu and Angrist, 2000) and the Metropolitan Statistical Area level (Moretti, 2004) could be partly explained by the rate of spatial decline in the importance of proximity to college for spillovers.

workers cannot be captured by the fixed boundaries of local areas. This is connected to decreasing search costs triggered by the availability of modern technologies. Online job search potentially enlarges labor markets as workers can search for distant jobs at very low cost. Finally, geographical areas are identical for any type of worker while the local availability of jobs and preferences towards job mobility can be very heterogeneous across subgroups of the working force.

In this paper, I propose a new and flexible approach to endogenously determine the size and shape of labor markets. Rather than depending on predefined geographical boundaries, endogenous labor markets are revealed by common patterns in the observed worker flows between firms. The approach is based on a network view of the labor market, where firms are linked to each other through worker transitions. Building on the universe of job-to-job transitions in the economy, I construct a job mobility network that reflects actual market interactions between firms. In particular, the firms in the economy constitute the nodes in the network and are connected by job-to-job transitions which generate directed and weighted links.[2] I partition this job mobility network into separate markets adapting a model from the literature on statistical network analysis. The basic idea in this novel approach is that two firms are in the same labor market if they have similar probabilities to link to the rest of the network and not because they are located in the same geographical area. This captures the possibility that – in addition to observed characteristics such as region and industry – labor markets are determined by unobserved factors. Consider, for instance, a market that contains firms which are employing computer scientists with expertise in a specific programming language, a market for jobs in an elite political class that can only be accessed by graduates of certain schools, or even a market that is characterized by a common dress code.[3]

The separation into endogenous labor markets is based on the stochastic block model (SBM) (Holland, Laskey, and Leinhardt, 1983; Karrer and Newman, 2011) which is the workhorse model for the detection of communities in the literature on network analysis. In my adaption of the SBM, firms are characterized by two sources of unobserved heterogeneity. First, they have an individual propensity to attract and release workers that captures firm-level differences in productivity and turnover. Second, they operate on separate, unobserved labor markets. Firms in the same labor market are characterized by common unobserved latent factors that determine their probability to link to each other. Worker transitions between firms are governed by the interplay of firm-level and market-level characteristics.

The SBM is related to a gravity-type equation where interactions between two agents are

[2] In this definition, links connect employers who draw on the same kind of skills. Moreover, the links entail information flows and spillover effects. The importance of job mobility for firm productivity and agglomeration economics is emphasized by a growing body of empirical research (Balsvik, 2011; Dasgupta, 2012; Poole, 2013; Stoyanov and Zubanov, 2012, 2014; Serafinelli, 2014).

[3] Indeed, the British Social Mobility Commission has recently identified obstacles to enter jobs in British investment banks that preclude market entrance of individuals who do not know the common code of conduct or dress code in some institutions.

determined by individual characteristics and a measure of distance. In the spirit of Eaton and Kortum (2002) and Cortes and Gallipoli (forthcoming), the SBM can therefore be micro-founded by a simple model of firm choice where job transitions are determined by the costs of switching. The main goal of the empirical analysis is then to identify sets of firms that have a similar structure of switching costs.

My empirical analysis is based on administrative records from the Austrian Social Security Database (ASSD) which provides detailed matched employer-employee data for all private sector employees in Austria. Building on the universe of job-to-job transitions from 1975 to 2005, I compute a large and detailed job mobility network and analyze the resulting network structure. The job mobility network comprises about 930,000 job-to-job transitions and more than 95,000 firms. Given this observed job mobility network, I estimate the SBM by maximum likelihood via Markov Chain Monte Carlo methods in order to assign firms to endogenous labor markets conditional on their individual propensity to attract workers.

In an extensive descriptive analysis, I examine the characteristics of my endogenous labor markets. Compared to local labor markets, endogenous labor markets are more self-contained. I find higher shares of job transitions within endogenous labor markets than within geographical entities of the same size. The estimated markets nevertheless reveal a "local" structure of job transitions as firms within the same endogenous market are spatially clustered. The resulting geographical structure of endogenous labor markets however deviates from administrative borders in three important ways. First, there are several largely unconnected endogenous markets located within the same geographical region. The separation of these markets can be partly rationalized by differences in the industry composition and wage distribution. Second, some endogenous labor markets are scattered across several regions and contain very distant firms. Geographically dispersed endogenous markets tend to be relatively more specialized in particular industries than geographically concentrated markets. In general, however, endogenous labor markets are not particularly concentrated within specific industries. Workers regularly switch industries and firms hire from a variety of different occupations.[4] Third, the geographical structure of endogenous labor markets varies over time. In particular, the average spatial distance between firms in the same market has increased by about 30% between the early 1980s and 2000s. In contrast, the industry dispersion within endogenous labor markets has slightly decreased over time.

The flexibility of the SBM further allows to examine heterogeneity in the scope of endogenous labor markets for various subgroups of the working force. I separately analyze job transitions by gender, age group, nationality, and skill-level. Most importantly, endogenous labor markets differ substantially between high- and low-skilled individuals. On average, the spa-

[4] This is consistent with evidence from other countries. For instance, Bjelland, Fallick, Haltiwanger, and McEntarfer (2011) show that more than 60% of job-to-job transitions in the US are reallocations across the 11 NAICS super-sectors.

tial distance between firms within endogenous labor markets for high-skilled individuals is about 1.3 times bigger than for low-skilled individuals. In contrast, markets for high-skilled workers are more concentrated within particular industries than markets for low-skilled workers. While endogenous labor markets of female workers were more specialized in particular industries in the 1980s, these differences vanished in the late 1990s.

My novel approach for the estimation of endogenous labor markets can provide new answers to a range of important economic questions. In the present paper, I utilize the model to analyze the impact of local labor demand shocks and global trade shocks on employment and worker mobility. In particular, I demonstrate that endogenous labor markets estimated in the period prior to the shock can explain and predict worker flows in response to the break-down of the Austrian steel industry at the end of the 1980s and to the increasing exposure of manufacturing industries to trade with China and Eastern Europe. In the first application, I examine employment spillovers following a series of unexpected mass layoffs in the Austrian steel industry in 1986. The break-down of the steel industry caused adverse effects on employment in firms from the same endogenous labor market, both in the region of the shock as well as in distant regions. At the same time, employment in firms from the same region but different endogenous labor markets remained unaffected. Transition probabilities between endogenous labor markets were also affected by the shock. In particular, there were less transitions into the affected market and more transitions out of this market. Importantly, the change in worker transitions was proportional to the predicted pre-shock transition probability. In the second application, I exploit the astonishing rise in trade with China and Eastern Europe in the past decades and quantify the relative importance of different margins of mobility adjustments. The negative impact of import competition from eastern countries on wages and employment of manufacturing workers can be partly offset by job mobility. In an analysis that follows the identification strategies of Autor, Dorn, Hanson, and Song (2014) and Dauth, Findeisen, and Suedekum (2016), I show that the endogenous labor markets estimated in the period prior to the shock can accurately predict job mobility responses to trade shocks. In contrast, markets based on geographical areas fail to explain important parts of the worker movements.

The remainder of the paper is organized as follows. I discuss the related literature in Section 1.2. Section 1.3 describes the data, the definition of the job mobility network, and aggregate network characteristics. In Section 1.4, I explain the stochastic block model and the estimation strategy in detail. I provide a descriptive analysis of the endogenous labor markets and evidence for worker heterogeneity in section 1.5. Section 1.6 examines spillover effects and mobility responses to local demand shocks and global trade shocks. Section 1.7 concludes.

1.2 Related Literature

My proposed new model contributes to a relatively new literature on alternative definitions of local labor markets. In a recent contribution to this literature, Manning and Petrongolo (forthcoming) endogenously determine the size of a spatial but flexible concept of local labor markets by optimal job search strategies of unemployed individuals. They find relatively narrow local markets as workers' search effort is sharply declining with distance to a vacancy. Lalive et al. (2015) use particular information on the characteristics of vacancies to predict whether two unemployed individuals would apply for the same job and hence be in the same market. Lechner, Wunsch, and Scioch (2013) exploit information on firm and worker location to determine hiring regions of workers. Commonly, however, the definition of labor markets in these papers is based exclusively on observable characteristics. My approach contributes to this literature by explicitly incorporating unobserved determinants of labor markets. I provide evidence for the importance of these unobserved determinants in the analysis of spillover effects and mobility responses to economic shocks.

My approach is also related to Schmutte (2014) who employs computer-based community detection algorithms to determine the boundaries of job mobility using data from the Panel Study of Income Dynamics (PSID). He finds four large segments of the labor market that do not coincide with industry, occupation, or education categories.

The method proposed in this paper adds to a rapidly growing literature that studies worker flows across firms in order to gain insight into the quality and preferences of workers and firms. Card, Heining, and Kline (2013) and subsequent papers revived the interest in the estimation of wage decompositions in the tradition of Abowd, Kramarz, and Margolis (1999). Similar to the job mobility network in my paper, the identification of worker and firm fixed effects in these studies is based on the set of firms that are connected by worker mobility. In a recent contribution, Sorkin (2015) exploits worker flows across firms to reveal preferences for non-wage characteristics of firms and compensating differentials. My paper complements these approaches in detecting common unobserved market-level characteristics of firms that are revealed by common patterns in worker flows.[5]

As shown in the application of my model to large economic shocks, estimating endogenous labor markets based on the SBM generates new insights into mobility adjustments to local shocks. The paper therefore contributes to an important literature that analyzes and estimates the incidence of shocks to local labor demand and supply (see, e.g., Blanchard and Katz., 1992; Bound and Holzer, 2000; Notowidigdo, 2013). While existing papers take the size of the local labor markets as given using predefined regions, endogenizing labor markets allows for various sources of heterogeneity in these effects. This also holds true for related

[5]The goal of endogenously grouping firms is also pursued in recent contributions by Bonhomme and Manresa (2015) and Bonhomme, Lamadon, and Manresa (2016). In contrast to my paper, the assignment of firms to groups is based on similarities in outcomes such as the distribution of wages in these papers.

studies that examine spillover effects of positive or negative shocks to local economies (Greenstone, Hornbeck, and Moretti, 2010; Busso, Gregory, and Kline, 2013; Gathmann, Helm, and Schönberg, 2016). In two recent papers, Cestone, Fumagalli, Kramarz, and Pica (2016) and Giroud and Mueller (2017) show that local economic shocks are absorbed through worker flows within internal labor markets that consist of firms affiliated to the larger groups. The empirical evidence in this paper confirms the view that economic ties are more relevant for the transmission of economic shocks than geographical proximity.

Finally, my paper adds to a very recent literature that tries to incorporate methods and insights from machine-learning into economics. To the best of my knowledge, it is the first application of the well-established stochastic block model from the statistics literature to an economic question. The SBM (Holland et al., 1983; Karrer and Newman, 2011) is the workhorse model in the literature on statistical networks to partition networks into groups based on the observed linkages. This task has also been coined community detection. In the original stochastic block model of Holland et al. (1983), all nodes in the same community behave stochastically equivalent and have the same probability distribution of links. The modified version of Karrer and Newman (2011) allows for heterogeneity within communities by preserving the observed distribution of connections. This is achieved by including node-specific fixed degree parameters and relates the approach to a recent literature of network formation with unobserved individual heterogeneity (Graham, forthcoming; Dzemski, 2014). Community detection has a long tradition in physics and computer sciences and has given rise to a variety of methods and algorithms (for an overview see Fortunato, 2010). The methods can be roughly classified into greedy ad-hoc algorithms such as hierarchical clustering (e.g. Clauset, Newman, and Moore, 2004), algorithms that optimize global criteria over all possible network partitions (e.g. the modularity score of Newman and Girvan, 2004), and model-based methods. In this paper, I consider a model-based approach, which makes the underlying assumptions and structure explicit.[6] This novel approach of unsupervised machine learning is broadly applicable to other economic contexts as for instance the analysis of supplier relationships among firms, trade networks, and others.

1.3 Data

The empirical analysis is based on administrative records for the universe of private sector employment in Austria. The matched employer-employee data from the Austrian Social Security Database (ASSD, see Zweimüller, Winter-Ebmer, Lalive, Kuhn, Wuellrich, Ruf, and Büchi, 2009) provides detailed daily information on employment and unemployment spells as well as on other social security related states such as sickness, retirement, or maternity leave since 1972. Each individual employment spell is linked to an employer identifier and some

[6] For an application of modularity score maximization in the context of job mobility see Schmutte (2014).

firm-level information.[7] Moreover, annual earnings are provided for each worker-firm combination.

To define the job mobility network, I extract all job-to-job transitions from the ASSD that occurred between 1975 and 2005 and satisfy the following criteria: First, a change of employer is classified as a job-to-job transition if there are at most 30 days of non-employment in between two consecutive employment spells. Second, the sample is restricted to transitions where workers had a minimum tenure of one year in both their old and new job. This allows me to examine only relatively stable relationships that are not prone to seasonal fluctuations.[8] Third, the sample is restricted to transitions between firms with five or more employees. Fourth, I exclude transitions of apprentices, marginal, and contract workers. Finally, spurious transitions due to firm renaming, spin-offs, and takeovers are excluded using the worker flow approach detailed in Fink et al. (2010).

Table 1.1 shows summary statistics of these job-to-job transitions in Austria. The first column refers to all job-to-job transitions that occurred in the entire observation period between 1975 and 2005, in the other columns I split the sample into job-to-job transitions from shorter periods in order to track developments over time. In total, there are more than 930,000 job-to-job transitions. Over time, the number of transition exhibits an upward trend.

The share of job-to-job transitions by female workers amounts to 41% and increases over time. About 96% of all transitions are experienced by Austrians with a slight decrease from 97% to 95%. The average age at the beginning of the new spell fluctuates between 31 and 35 years.

When switching their job, workers on average engage in more stable relationships. The average spell duration at the old firm amounts to 1550 days, while the new job last for 2350 days on average. The average time in between two spells decreased from 4.4 days in the late 1970s to 3.5 days in the early 2000s.

Regional mobility is relatively low as about 75-80% of the job switchers transit to a firm in the same state and 65% remain in their NUTS-3 region.[9] In contrast, mobility across industries is much higher as about two thirds of the job-to-job transitions occur between firms that are affiliated to different NACE 2-digit industries. This finding is consistent with recent evidence from the US (Bjelland et al., 2011) where about 60% of job switches are reallocations across the 11 broad NAICS super-sectors. Over time, however, there are opposite trends in

[7] I use the terms employer identifier and firm interchangeably. Fink, Kalkbrenner, Weber, and Zulehner (2010) compare the distribution of employers in the ASSD with official firm registers from Statistics Austria and find only negligible differences. Hence, they conclude that multi-establishment firms are not an important component in the Austrian market.

[8] A substantial part of the Austrian economy is characterized by seasonal sectors such as construction and tourism. An alternative version of the model that includes jobs with minimum spell duration above one month does not change the results substantially but generates additional noise.

[9] Austria consists of 9 very heterogeneous states. The "Nomenclature of territorial units for statistics" classification of Eurostat counts 35 NUTS-3 regions in Austria, see also the maps in figure 1.5. NUTS-3 regions in Austria are aggregations of several municipalities.

Table 1.1: Summary Statistics of Job-to-Job Transitions

	1975-2005	1975-1980	1980-1985	1985-1990	1990-1995	1995-2000	2000-2005
Demographics							
share of females	0.41	0.37	0.37	0.41	0.42	0.41	0.42
	(0.492)	(0.482)	(0.489)	(0.491)	(0.495)	(0.492)	(0.494)
share of Austrians	0.96	0.97	0.97	0.97	0.94	0.94	0.95
	(0.201)	(0.180)	(0.173)	(0.174)	(0.233)	(0.234)	(0.222)
avg. age	33.0	33.3	31.9	31.0	32.0	33.4	34.7
	(9.75)	(10.69)	(10.16)	(9.50)	(9.34)	(9.12)	(9.28)
Avg. duration (in days) of							
spell at old firm	1545.9	1158.0	1429.5	1491.4	1509.6	1583.5	1625.4
	(1445.52)	(659.41)	(1075.05)	(1317.09)	(1462.83)	(1567.81)	(1657.43)
spell at new firm	2354.0	2650.2	2613.0	2389.4	2237.3	1997.0	1596.4
	(2153.34)	(2687.41)	(2565.27)	(2234.63)	(1890.99)	(1443.00)	(907.22)
intermission between spells	4.0	4.4	4.2	4.0	4.1	3.7	3.5
	(5.98)	(6.33)	(6.22)	(5.92)	(6.06)	(5.70)	(5.56)
Share of workers staying in the same							
state	0.78	0.79	0.78	0.78	0.79	0.78	0.74
	(0.417)	(0.409)	(0.414)	(0.411)	(0.407)	(0.417)	(0.440)
NUTS-3 region	0.64	0.66	0.65	0.66	0.66	0.66	0.63
	(0.479)	(0.473)	(0.476)	(0.474)	(0.472)	(0.475)	(0.484)
2-digit industry	0.22	0.21	0.21	0.20	0.23	0.26	0.28
	(0.417)	(0.410)	(0.404)	(0.400)	(0.422)	(0.441)	(0.447)
occupation (white- and blue-collar)	0.86	0.87	0.85	0.85	0.86	0.87	0.88
	(0.346)	(0.341)	(0.355)	(0.361)	(0.351)	(0.332)	(0.320)
Share of workers with							
wage increase	0.61	0.65	0.63	0.64	0.63	0.60	0.60
	(0.487)	(0.476)	(0.484)	(0.481)	(0.483)	(0.492)	(0.490)
increase in firm fixed effect	0.74	0.89	0.86	0.83	0.78	0.72	0.62
	(0.441)	(0.31)	(0.35)	(0.379)	(0.415)	(0.451)	(0.485)
Number of transitions	930,027	258,837	204,832	234,580	273,099	263,006	281,880

Note: Standard deviations in parenthesis. This table reports summary statistics for job-to-job transitions in the specified period. A change of employer is classified as a job-to-job transition if there are at most 30 days of intermission in between two consecutive employment spells and tenure at both employers exceeds one year. Transitions between small firms with less than 5 employees and transitions of apprentices and marginal workers are excluded.

mobility across regions and industries. While regional persistence has slightly decreased in the beginning of the 2000s, there is a clear upward trend in the share of workers that remain in their industry starting from the early 1990s. The vast majority of workers remain in their broad occupation (white- or blue-collar workers) with a slight increase in later periods.

In general, most workers move up the job ladder when switching their employer. For more than 60% of the workers the transition is associated with a wage raise. Moreover, 74% of transitions lead workers to a firm with a higher firm fixed effect as measured by a wage decomposition as in Abowd et al. (1999).[10] Interestingly, however, this share deteriorates sharply over time from 90% in the late 70s to only 62% in the early 2000s.

The main idea behind defining the job mobility network is that job-to-job transitions establish links between the two firms involved in the transition. Figure 1.1 depicts the concept of link formation. When workers flow between firms i and j (Figure 1.1a), directed and weighted links are established between those firms (Figure 1.1b). In particular, the link from i to j is stronger the more transitions occur in this direction during the sample period.

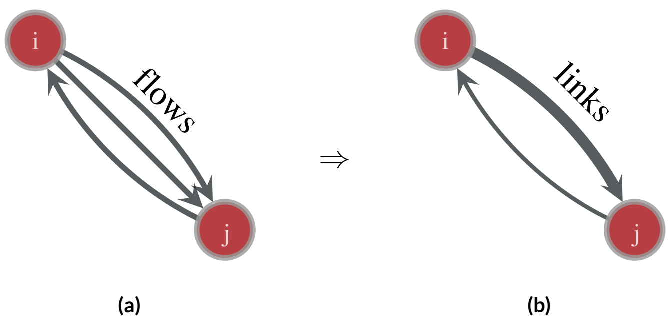


Figure 1.1: Link Definition Based on Job-to-job Transition

Applying this procedure to the universe of job-to-job transitions in Austria results in a large and very detailed job mobility network. Formally, the job mobility network $G = \{V, E\}$ consists of a set of N nodes $V = \{1, 2, \ldots, N\}$, i.e., the firms in the economy, and a set of links E, i.e., the job-to-job transitions.[11] An $N \times N$ adjacency matrix A indicates which firms are linked and how strong the ties are. Particularly, A_{ij} denotes the number of job-to-job transitions from firm i to firm j within the sample period.[12]

[10] Firm fixed effects and worker fixed effects are obtained from estimating the following equation for log earnings

$$y_{wt} = \alpha_w + \Psi_{J(w,t)} + x'_{wt}\beta + \varepsilon_{wt},$$

where y_{wt} are log earnings of worker w at time t, α_w is a worker fixed effect, $\Psi_{J(w,t)}$ is the firm fixed effect at firm j where worker w is employed at time t, x contains a set of covariates such as age and education, and ε is an error term. This decomposition is known as the AKM decomposition (after Abowd et al., 1999) and will be used at several points in the paper.

[11] More precisely, V contains only the non-isolate firms in the economy, that is, firms, which are involved in at least one job-to-job transition during the sample period (see also appendix 1.A).

[12] A formal definition of the adjacency matrix is provided in Appendix 1.A. The job mobility network defined in this paper differs from the approach proposed in Schmutte (2014) in a number of important ways. First, it

There are two important economic aspects of link formation in the job mobility network. First, links connect firms that employ the same worker within a short period of time. This ensures homogeneity of closely connected firms as they are drawing on the same worker type and skill set. Second, links are potential channels for information flows and knowledge spillovers between firms. Recent theoretical and empirical work has pointed out the importance of job mobility for knowledge transfers and spillovers (see Balsvik, 2011; Dasgupta, 2012; Poole, 2013; Stoyanov and Zubanov, 2012, 2014; Serafinelli, 2014). These papers show that it matters to which firms a firm is connected as incoming workers from highly productive firms often generate positive productivity spillovers. Cestone et al. (2016) and Giroud and Mueller (2017) show that job mobility between tightly linked firms can also serve as an insurance mechanism within internal labor markets of larger corporate companies.

Table 1.2: Links and Nodes in the Job Mobility Networks

	1975-2005	1975-1980	1980-1985	1985-1990	1990-1995	1995-2000	2000-2005
# of nodes	95,237	54,080	51,406	54,615	59,215	59,444	61,068
# of links	930,027	258,837	204,832	234,580	273,099	263,006	281,880
# of components	755	1232	1549	1407	1265	1725	1889
in giant component							
% of nodes	0.98	0.95	0.93	0.94	0.95	0.94	0.93
% of links	1.00	0.99	0.99	0.99	0.99	0.99	0.99
average degree	19.84	9.98	8.42	9.00	9.59	9.31	9.75

Note: All measures correspond to the job mobility network sampled during the years indicated. The number of components counts all subgraphs of the network within which all firms are connected by some path, but not connected to the other subgraphs. Avg. degree measures the average number of incoming and outgoing connections per firm.

Table 1.2 provides an overview of the number of links and nodes in the job mobility network. Again, column 1 refers to the full network obtained from transitions between 1975 and 2005 while the other columns show dynamic developments between shorter periods. In total, there are more than 95,000 firms (nodes) in the network and about 930,000 transitions (links). Over time, these figures tend to rise, reflecting the increase in job mobility and the number of firms involved. The third row of Table 1.2 shows the number of connected components in the network. A connected component is a subgraph of the job mobility network within which all firms are connected by some path, but not connected to the other subgraphs. In the full network, there are 755 components. The vast majority of these components, however, contains only 2 firms while the largest connected component (the *giant component*) contains about 98% of the firms and almost all links (rows 4 and 5). The analysis in the remainder of the paper is therefore restricted to the giant component.[13]

corresponds to the one-mode employer projection graph of his realized mobility network but additionally allows for *directed* links that capture actual flows. Second, it distinguishes *direct* and *indirect* connections between firms while this is not possible in his approach where all firms a worker has worked for at any time are directly connected.

[13]This restriction is identical to limitations in the literature on AKM-type wage decompositions where

The last row of Table 1.2 displays the average degree in the network which denotes the average number of transitions per firm. On average, a firm is connected to 19.84 other firms in the full network. The average degree is naturally lower when considering shorter time periods and tends to slightly rise over time. The average degree however hides substantial heterogeneity within the network. Many firms in the job mobility network are involved only in a low number of job-to-job transitions while others have many connections and serve as "hubs" in the economy. In appendix 1.A, I provide a detailed discussion on various network characteristics and document the heterogeneity in the number of connections. In my model for the estimation of endogenous labor markets in the following section, I specifically address firm-level heterogeneity by including popularity parameters that guide the individual attractiveness of firms to workers.

1.4 Model

In this section, I present a novel method that allows to endogenously determine labor markets based on common patterns of worker flows observed in the job mobility network. First, I set up a model of job mobility that generates job-to-job transitions which are restricted by transition costs and search frictions. The model is based on a model of occupational flows in Cortes and Gallipoli (forthcoming) which in turn draws on the literature in international trade, particularly on Eaton and Kortum (2002). The key equation of the job mobility model is a gravity equation that relates flows between firms to (market-level) transition costs. In a second step, I use the Stochastic Block Model from the literature on network analysis in order to estimate the most likely assignment of firms to markets using this equation.

The model economy is populated by a finite set of firms indexed by $i \in \{1, \ldots, N\}$ and a continuum of workers of measure one indexed by ℓ. Workers differ in observable characteristics and their initial firm. A worker's firm at the beginning of the period is predetermined and given by i.

The potential payoff from switching to firm j for worker ℓ who is currently employed at firm i is denoted by

$$\varphi_j(\ell|i) = p_j f[X_\ell] \left(\frac{x_j(\ell)}{d_{ij}} \right) \tag{1.1}$$

where p_j is a single index subsuming firm-level factors that affect all workers at firm j (i.e., the general attractiveness of j), X_ℓ is a vector of individual characteristics that change returns for worker ℓ in all firms (as for instance general human capital of worker ℓ), $x_j(\ell)$ is a worker specific match-quality shock that measures how well worker ℓ is matched with firm j in terms of productivity and preferences, and d_{ij} represents the costs of switching from firm i to j.

worker and firm fixed effects are only separately identified within connected sets of firms that are linked by worker mobility (Abowd, Creecy, and Kramarz, 2002).

I assume that the costs of transitioning between firms are not firm-pair specific but are determined on the market level. Specifically, each firm operates on one of k different markets in the economy. An $N \times 1$ vector z denotes the assignment of firms to markets with $z_i \in \{1, \ldots, k\}$. Formally, $d_{ij} = d_{z_i z_j}$ and two firms i and j are in the same market, i.e., $z_i = z_j$ if the utility costs of moving to other firms in the economy are identical. I normalize costs such that moves within the same market have a cost of one, $d_{uu} = 1$, while switches across markets are associated with positive costs, $d_{uv} > 0 \, \forall v \neq u$. These utility costs could represent pure moving costs that depend on spatial distance.[14] However, these costs could also represent time and efficiency costs associated with adapting to the new firm, skill transferability between industries and occupations, or other – unobserved – components. The main idea of the approach is to determine the assignments of firms to markets endogenously, that is use observed transitions to identify firms that compete for the same worker by a revealed preference argument.

Match-quality is drawn from a Fréchet distribution

$$x_j \sim F_j(x) = \exp(-T_j x^{-\theta}) \tag{1.2}$$

where $T - j$ is a firm-specific location parameter and θ governs the dispersion of the shock. At the beginning of the period, workers receive the opportunity to examine outside options with arrival rate λ. If they get this opportunity, they sample match-qualities, drawing a value for each firm. They then compare potential payoffs based on the realized draws and decide whether to switch and where to go.

From the distributional assumption of match-quality it follows that the probability of a switch from firm i to firm j is

$$\pi_{ij}(\ell) = \underbrace{\lambda}_{\text{switching opportunity}} \times \underbrace{\frac{T_j d_{z_i z_j}^{-\theta} p_j^{\theta}}{\sum_{s=1}^{N} T_s d_{z_i z_s}^{-\theta} p_s^{\theta}}}_{j \text{ offers the highest payoff of all firms for a worker who starts in } i}. \tag{1.3}$$

This result is borrowed from the literature on international trade (Eaton and Kortum, 2002) and derived in more detail in Appendix 1.B. Notably, the switching probability does not dependent on individual worker level characteristics.

Normalizing by the probability that origin firm i offers the highest payoff, we get an expression that relates worker flows to a set of firm-level characteristics and the transition costs.

$$a_{ij} = \lambda \times \frac{T_j d_{z_i z_j}^{-\theta} p_j^{\theta}}{T_i p_i^{\theta}} \tag{1.4}$$

According to the assumption of Poisson arrivals, the number of job-to-job transitions be-

[14] If markets are purely determined by geographical regions, this is consistent with the common assumption that moves within regions are costless while workers have to pay utility costs to move across regions.

tween any two firms i and j within a certain time period is an independent draw from the Poisson distribution

$$A_{ij} \overset{ind.}{\sim} Pois(\gamma_i^- \gamma_j^+ M_{z_i z_j}), \tag{1.5}$$

where γ_i^- summarizes firm-level characteristics of the firm of origin (i.e., T_i and p_i) and γ_j^+ summarizes destination firm characteristics (T_j and p_j). The $k \times k$ matrix M captures (the inverse of) the common cost component of transition probabilities within and between markets where the typical element M_{uv} indicates how likely a firm in market u experiences a job-to-job transition of one of its workers to a firm in market v.

This implies that the expected number of transitions from i to j, $E[A_{ij}] = \gamma_i^- \gamma_j^+ M_{z_i z_j}$, is increasing in the propensity of i to loose workers, the propensity of j to attract workers, and on the inverse of the transition costs between the markets of i and j.

In the data, the labor market assignments of firms are unobserved. Hence, the primary goal is to estimate these assignments given the observed job mobility network G described in section 1.3. For a given number of markets, k, the expression in equation (1.5) corresponds to the key equation of the Stochastic Block Model (SBM) from the literature on stochastic network formation.[15] Here we interpret the number of transitions between i and j as a weighted link between the two firms. The SBM can be summarized in the following likelihood function:

$$\begin{aligned}\mathcal{L}(G|M, z, \gamma) &= \prod_{i,j} Pr(i \to j|M, z, \gamma) \\ &= \prod_{i \neq j} \text{Poisson}(\gamma_i^- \gamma_j^+ M_{z_i z_j}) \\ &= \prod_{i \neq j} \frac{(\gamma_i^- \gamma_j^+ M_{z_i z_j})^{A_{ij}}}{A_{ij}!} \exp(-\gamma_i^- \gamma_j^+ M_{z_i z_j}).\end{aligned} \tag{1.6}$$

The product is taken over all combinations of i and j while self-loops (job-to-job transition of a firm to itself) are not allowed in the model. In order to identify the popularity parameters, I normalize the sum of all incoming and outgoing links in a market to one, $\sum_i \gamma_i^+ 1\{z_i = u\} = 1$ and $\sum_i \gamma_i^- 1\{z_i = u\} = 1$ for each market u. Imposing these constraints, the likelihood can

[15] The original stochastic block model (SBM) of Holland et al. (1983) and Wang and Wong (1987) defines a probability distribution over networks G, $Pr(G|z, M)$ that is guided only by the parameters z and M. The underlying assumption is that nodes within a group are stochastically equivalent, that is, all nodes from group u have the same independent probability of linking to a node from group v. Hence, the SBM does not allow for degree heterogeneity within groups. As there is typically ample variation in the connectedness of nodes in empirical networks (compare also figure 1.A14 for the present case), it is important to account for this kind of heterogeneity. Furthermore, in the original SBM the link variables are independent Bernoulli random variables. Simulation evidence in Zhao, Levina, and Zhu (2012) however shows that the difference between Bernoulli and Poisson is negligible especially with many nodes and small interaction probabilities.

be simplified to

$$\mathcal{L}(G|M,z,\gamma) = \Big(\prod_{i\neq j} A_{ij}!\Big)^{-1} \prod_i (\gamma_i^-)^{d_i^-}(\gamma_i^+)^{d_i^+} \prod_{u,v} M_{uv}^{E_{uv}} \exp\big(-M_{uv}\big) \tag{1.7}$$

where $d_i^- = \sum_j A_{ij}$ and $d_i^+ = \sum_j A_{ji}$ denote out- and indegree of firm i, and $E_{uv} = \sum_{ij} A_{ij} 1\{z_i = u\} 1\{z_j = v\}$ is the total number of links between firms in markets u and v.

Given the observed job mobility network network G, the model parameters can be estimated in a two-step procedure. In the first step, maximum likelihood estimators for M, γ^+, and γ^- conditional on a partition z are derived from the logarithm of equation (1.7). In particular, taking derivatives of the log-likelihood under the identification constraint yields

$$\hat{\gamma}_i^+ = \frac{d_i^+}{\delta_{z_i}^+}, \quad \hat{\gamma}_i^- = \frac{d_i^-}{\delta_{z_i}^-}, \quad \text{and} \quad \hat{M}_{uv} = E_{uv}, \tag{1.8}$$

where $\delta_u = \sum_{i:z_i=u} d_i$ denotes the sum of degrees in group u. These maximum likelihood estimators are very intuitive as relative popularity is measured by the relative number of connections and transition probabilities are measured by the number of observed transitions. Substituting the estimators in equation (1.8) into the (log-)likelihood and neglecting terms that do not depend on the model parameters considerably simplifies the expression to

$$\ln \mathcal{L}(G|z) = \sum_{u,v} E_{uv} \ln \frac{E_{uv}}{\delta_u^+ \delta_v^-}, \tag{1.9}$$

which depends only on the counts induced by the choice of the partition z.

In the second step, the log-likelihood is maximized by choosing the partition of firms into markets z which maximizes (1.9). Since it is not feasible to evaluate all possible combinations of firms and markets, the empirical analysis relies on computational approximations via a Markov-chain Monte-Carlo (MCMC) algorithm. In particular, the market assignments of the firms are modified in a random fashion and each move is accepted or rejected depending on the change in the likelihood (for details of the algorithm see Peixoto, 2014a). The estimation is repeated for different starting partitions in order to avoid lock-in at local maxima.

A generalized consistency framework for community detection using the SBM is provided by Bickel and Chen (2009) and (including the popularity parameters) by Zhao et al. (2012).

The remaining issue pertains to the choice of the number of markets k which has been treated as fixed so far. This parameter guides the "size" of the model as a larger k implies more parameters in the transition matrix M. It can also be used to "zoom" into or out of the economy in order to analyze different levels of market aggregation. However, a tradeoff between more flexible models and the threat of over-fitting arises. In the empirical analysis, I estimate the SBM for various choices of k and evaluate the different fits using the modularity score of Newman and Girvan (2004).[16] The modularity score measures how well a network decomposes into self-contained communities. A high score indicates dense connections between firms within markets but only sparse connections between firms from different markets.[17] Figure 1.2 displays the modularity score for varying k in the job mobility network for the years 1975-2005. There is a clear peak at $k = 9$ groups, indicating that a SBM with nine markets is best suited to describe the structure of labor markets in Austria.

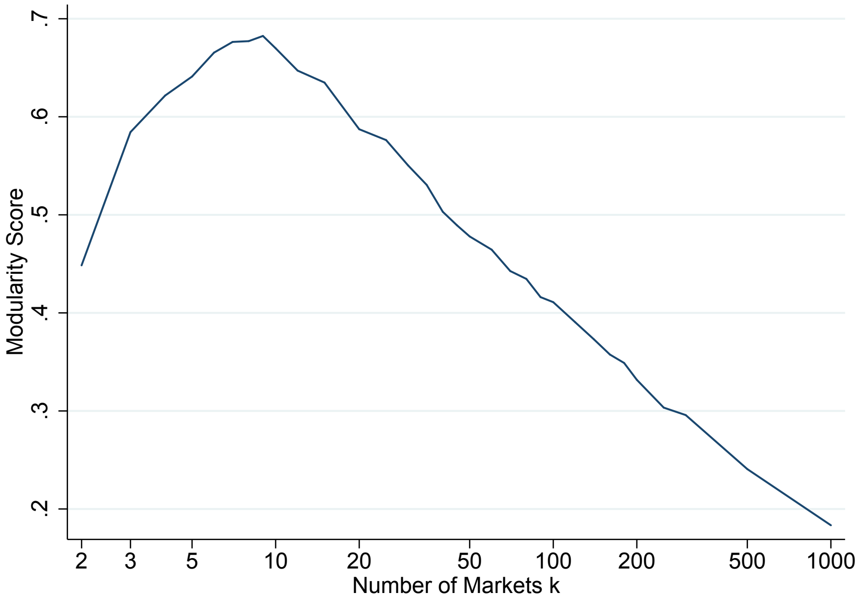


Figure 1.2: Modularity Score for Varying Number of Groups k in the Job Mobility Network (1975-2005)

1.5 Descriptive Analysis of Endogenous Labor Markets

In this section, I present and discuss the endogenous labor markets that arise from partitioning the job mobility network among Austrian firms via the stochastic block model (SBM). I start by comparing endogenous labor markets to geographical characterizations of labor mar-

[16] In principle, regularization methods such as information criteria (BIC, AIC, etc.), minimum description length, or likelihood ratio tests, could guide the choice of k. Due to the complex asymptotic behavior of network models, however, traditional criteria are biased in many ways and finding corrections for model selection is an active strand of the statistical networks literature (see Yan, Shalizi, Jensen, Krzakala, Moore, Zdeborova, Zhang, and Zhu, 2014).

[17] In particular, the modularity score is defined as $Q = \frac{1}{2|E|} \sum_{ij} \left(A_{ij} - \frac{d_i d_j}{2|E|} \right) 1\{z_i = z_j\}$ where $|E|$ is the total number of transitions in the network. It compares the share of links within a market to the expected share in a model where all firms have the same number of links but links are generated uniformly at random (ignoring the market structure). This implies that the score is 0 if the markets have no explanatory power while a positive score indicates that there are more links within communities than expected under random link formation (cf. Jackson, 2008).

kets in section 1.5.1. In section 1.5.2, the analysis continues with a comparison of endogenous labor markets for several subgroups of the working force.

1.5.1 Endogenous versus Local Labor Markets

This section provides an extensive descriptive analysis of endogenous labor markets in Austria. Particularly, I compare the estimates from the SBM to local labor markets that are based on predefined geographical characteristics.

Self-Containedness

Figure 1.3 gives an overview of endogenous labor markets estimated based on the job mobility network for the time period from 1975 to 2005. Each circle represents one of the $k = 9$ sets of firms that has been assigned to the same market by the SBM.[18]

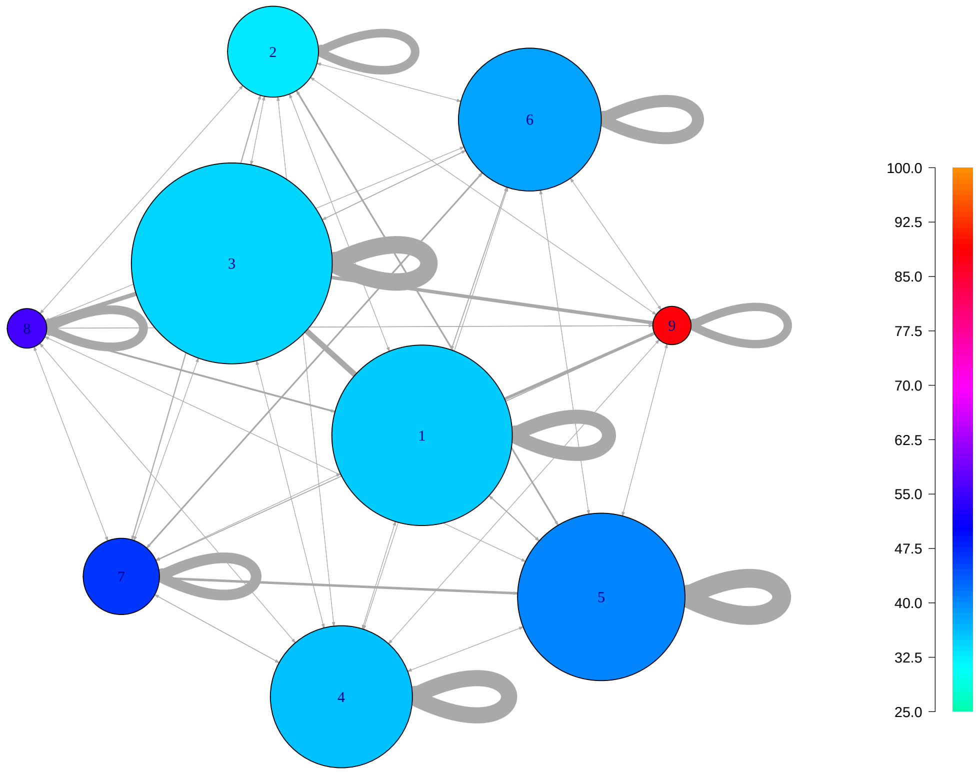


Figure 1.3: Estimated Block Network (node size $\propto$ market size, link width $\propto$ transition probability, colours $\propto$ avg. firm size)

The transition probabilities across and within markets are represented by the gray edges, which are thicker the more likely a transition is. Job-to-job transitions within markets are

[18] The numbering of the markets bears no particular meaning and only serves to label markets. The size of each circle is proportional to the number of firms in the respective market. The coloring of the circles represents the average firm size in the market. Evidently, firms in markets with many firms tend to be smaller on average than firms in markets with fewer firms.

much more probable than transitions between different markets. This clear segmentation can also be seen in Figure 1.4 which displays the estimated transition probabilities (normalized to sum up to one).[19] In total, 80% of all job-to-job transitions occur within endogenous labor markets.[20] Transitions between markets are much less likely. The closest connection is between markets 1 and 3 where 1.7% (from market 1 to market 3) and 1.9% (from 3 to 1) of all transitions occur.

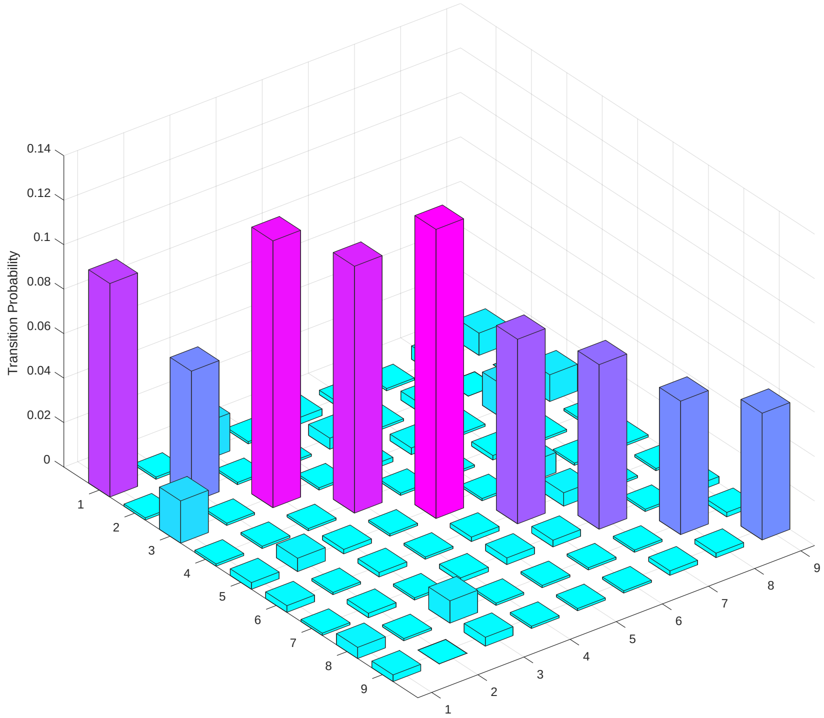


Figure 1.4: Estimated Transition Probabilities between Markets

A natural metric to evaluate how self-contained different definitions of labor markets are is the modularity score of Newman and Girvan (2004). It computes the share of transitions between firms *within* the same market among all transitions and normalizes it by the share of within market transitions that would be expected if links were generated uniformly at random (ignoring the market structure).[21] The modularity score therefore measures the explanatory power of the inherent market structure in excess of random link formation and a positive value indicates that there are more links within markets than expected. Table 1.3 displays the modularity score for endogenous labor markets with $k = 9$ and $k = 35$ as well as for markets defined by the 9 states, the 35 NUTS-3 regions, 2-digit industries, and state times 2-digit industry cells in Austria.

[19] The actual estimates are reported in Panel A of Table 1.A1 in the appendix.

[20] Recall that 77.5% of all transitions occurred within Federal states, 64.3% within Nuts-3 regions, and 22.4% within 2-digit industries.

[21] See also footnote 17.

Table 1.3: Modularity Scores for the SBM vs. Predefined Markets

	SBM with 9 markets	SBM with 35 markets	Federal States	NUTS-3 Regions	2-digit Industries	States × Industries
1975–2005	0.683	0.536	0.610	0.519	0.130	0.102
1975–1980	0.679	0.563	0.607	0.535	0.121	0.106
1980–1985	0.671	0.558	0.613	0.529	0.122	0.099
1985–1990	0.675	0.548	0.616	0.529	0.125	0.099
1990–1995	0.681	0.566	0.619	0.530	0.142	0.111
1995–2000	0.681	0.549	0.598	0.508	0.158	0.121
2000–2005	0.679	0.580	0.567	0.478	0.159	0.118

Note: This table reports modularity scores for labor markets estimated by the SBM (with k = 9 and k = 35) and defined by several observable characteristics. The modularity score compares the observed share of links within markets to the expected share in a model with the same degree distribution but random link formation. Higher values indicate more self-contained markets.

For the entire period from 1975 to 2005 and for each of the shorter sample periods, the SBM outperforms markets based on predefined geographical or industry characteristics. Endogenous labor markets with $k = 9$ have higher modularity scores than the 9 Austrian states and scores for endogenous markets with $k = 35$ exceed the scores for the 35 NUTS-3 regions in Austria. Not surprisingly, 2-digit industries and state by industry cells have much lower scores. Most importantly, the development over time indicates that the advantage of my novel method to determine endogenous labor markets grows with increasing mobility in the society.[22]

The Geography of Endogenous Labor Markets

The regional structure of the endogenous labor markets is illustrated in Figure 1.5. For each of the 9 markets, the figure presents a map with boundaries according to the NUTS-3 classification (Nomenclature of territorial units for statistics) of Eurostat. In each map, the 35 Austrian NUTS-3 regions are colored according to the share of firms from the relevant market within the respective region. There is clearly a *local* structure of labor markets in Austria as firms in the same endogenous market are geographically clustered. For each endogenous labor market, the vast majority of firms is concentrated within one Austrian state. The endogenous market structure however deviates from a classification that is solely based on geographical boundaries in two important aspects. First, firms in the same geographical area can be part of different endogenous labor markets. Second, sometimes distant firms from separate local labor markets are part of the same endogenous market.

[22] A second test of the performance of the SBM is provided in appendix 1.C. I conduct a Monte-Carlo simulation study with varying degrees of correlation between regions and true labor markets. The results demonstrate that even slight deviations from perfect correlation lead to the SBM outperforming regional characteristics.

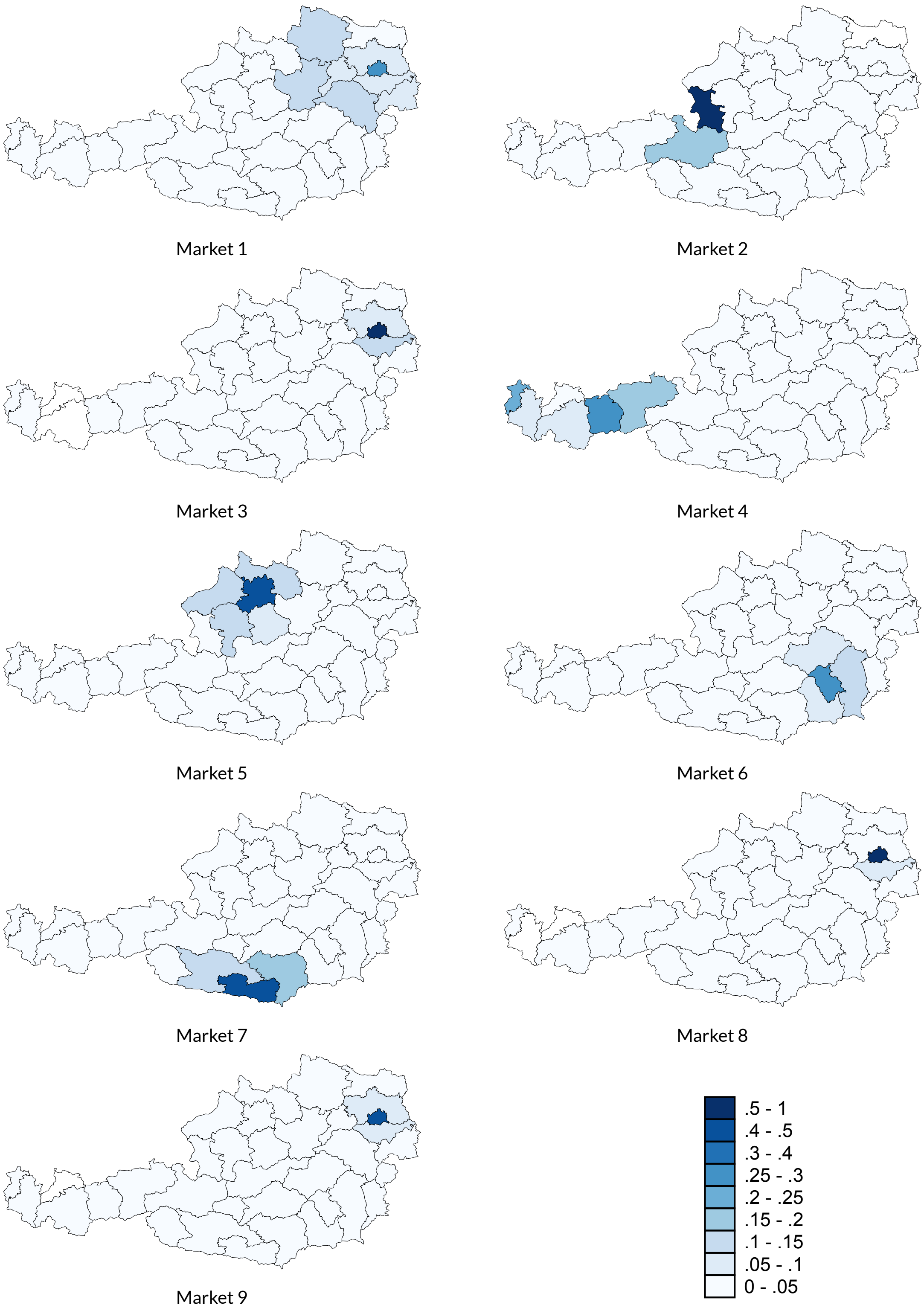


Figure 1.5: Share of Firms in NUTS-3 Regions for each Market (1975-2005)

Separate Endogenous Markets in the Same Area The maps in Figure 1.5 indicate several mostly separated endogenous labor markets within the same region. In particular, markets 3, 8, and 9 are all primarily located within Vienna. The estimated probability to switch between any of these markets is however less than 1.5%. Even within Vienna, the distribution of firms across different postal code areas is remarkably similar among these three markets (see the histograms in Figure 1.6).[23] Hence, firms in the same local labor market are located in different endogenous labor markets. It is therefore interesting to ask what distinguishes these endogenous markets from each other.

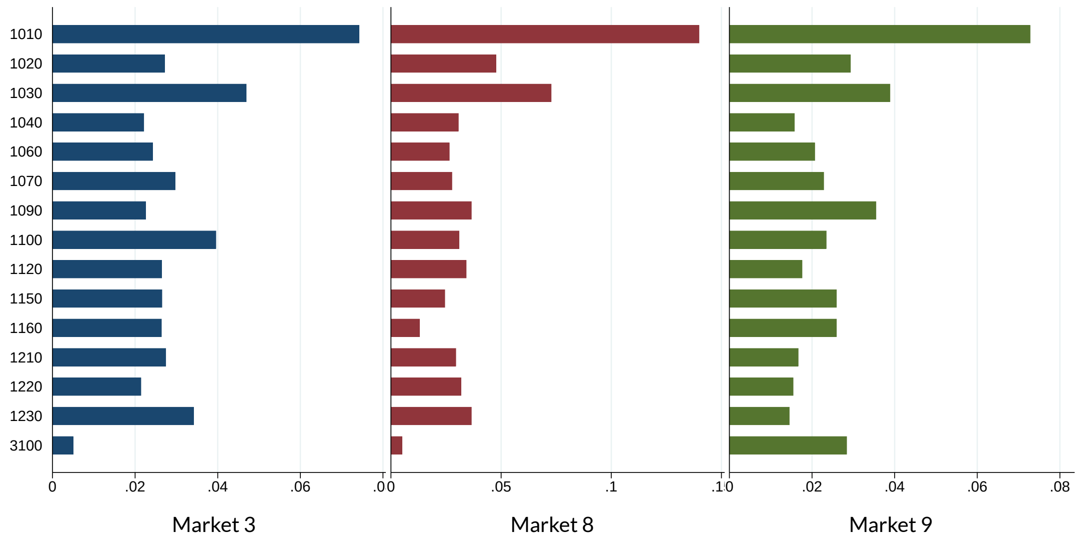


Figure 1.6: Histogram of Postal Code Areas for Markets in Vienna (1975-2005)

The histograms in Figure 1.7 illustrate the 2-digit industry composition of the three Viennese markets and reveal interesting differences between them. Despite the fact that substantial shares of firms in all three markets are affiliated to generic 2-digit industries such as Wholesale, Retail, and Construction, there are clear patterns of specialization. Many firms in market 3 are affiliated to manufacturing industries such as Food and Tobacco, Metal Products, Paper and Print, or the sale, maintenance, and repair of Motor Vehicles. Firms in market 8 are predominantly affiliated to Business Activities, Financial Services and Computer-related industries. Finally, firms from market 9 are specialized in Health, Public Administration, Lobbying, and Education.

Further important differences between the three Viennese markets can be found in terms of the wage structure. Figure 1.8 indicates that the distribution of firm fixed effects from an AKM wage decomposition is shifted to the right and more compressed in market 8 compared to markets 3 and 9. Not surprisingly, the market with a larger share of high-paying firms is the one specialized in Business and Financial Services.

[23]Note that Figure 1.6 (and Figure 1.7) only lists postal code areas (2-digit industries) where at least one of the markets has more than 2.5% of firms.

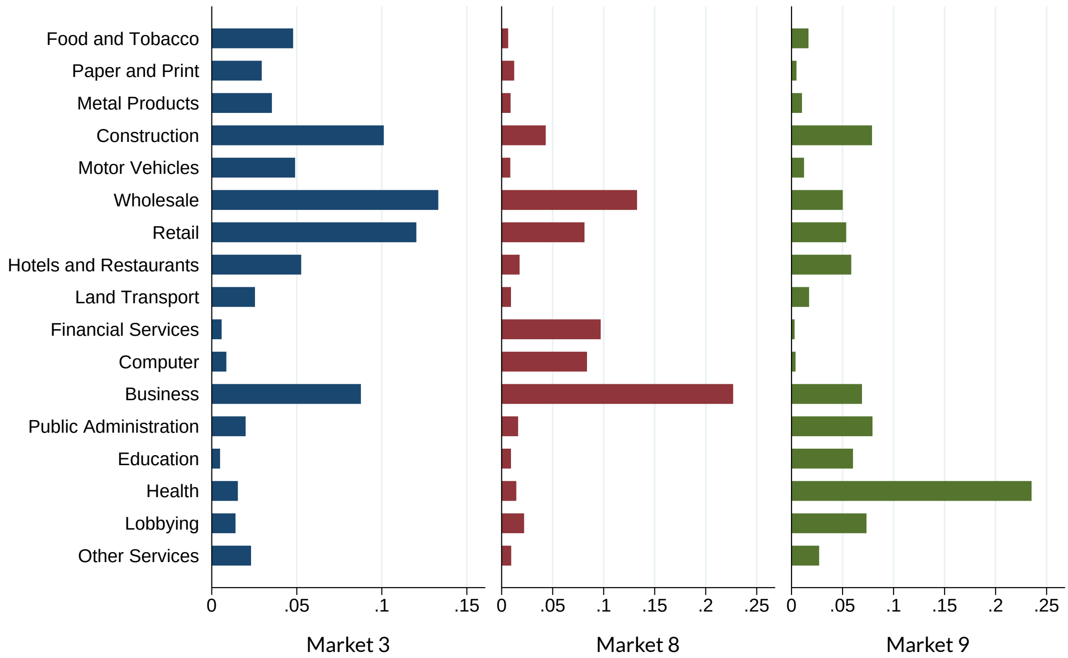


Figure 1.7: Histogram of 2-digit Industries for Markets in Vienna (1975-2005)

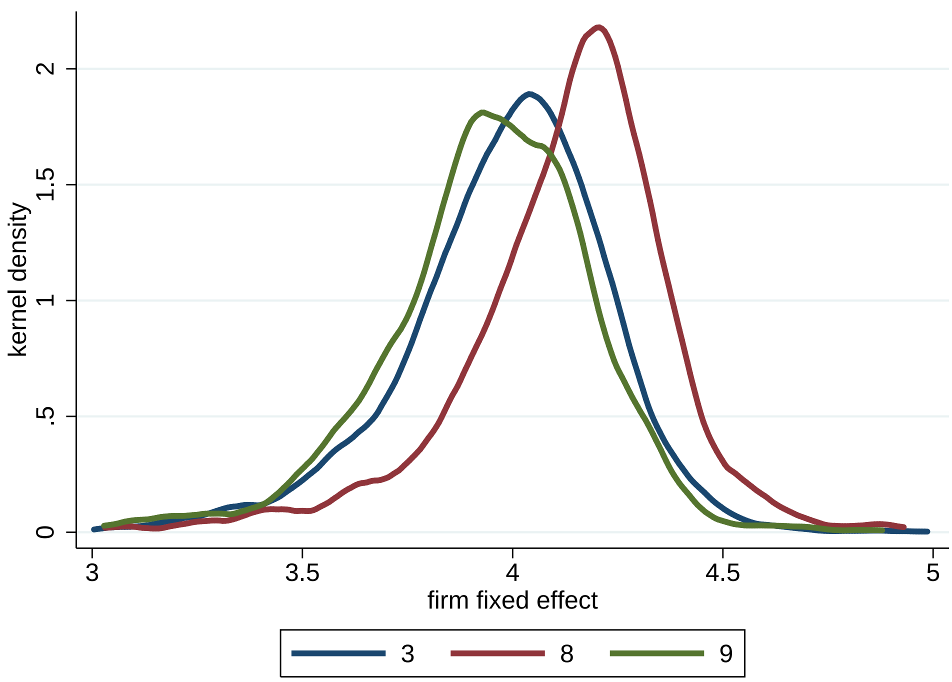


Figure 1.8: Kernel Density Estimates for the Distribution of Firm Fixed Effects for Markets in Vienna (1975-2005)

Distant Firms in the Same Endogenous Market Some endogenous labor markets are spread across a variety of regions and contain firms from a various states. While the industry affiliation of firms is in general not a good predictor of their assignment to endogenous labor markets, it becomes more important for markets that are spread out across the country.

For each of the 9 markets, Figure 1.9 provides a histogram of the broad sectoral composition. All markets consist of firms from a broad variety of industries. Evidently, sectors with a high degree of fluctuation such as construction, wholesale and retail, and hotels and restaurants are strongly represented in the sample of job-to-job transitions. There are however some markets with a stronger focus on particular sectors such as the dominance of professional services in market 8, or the health sector in market 9.

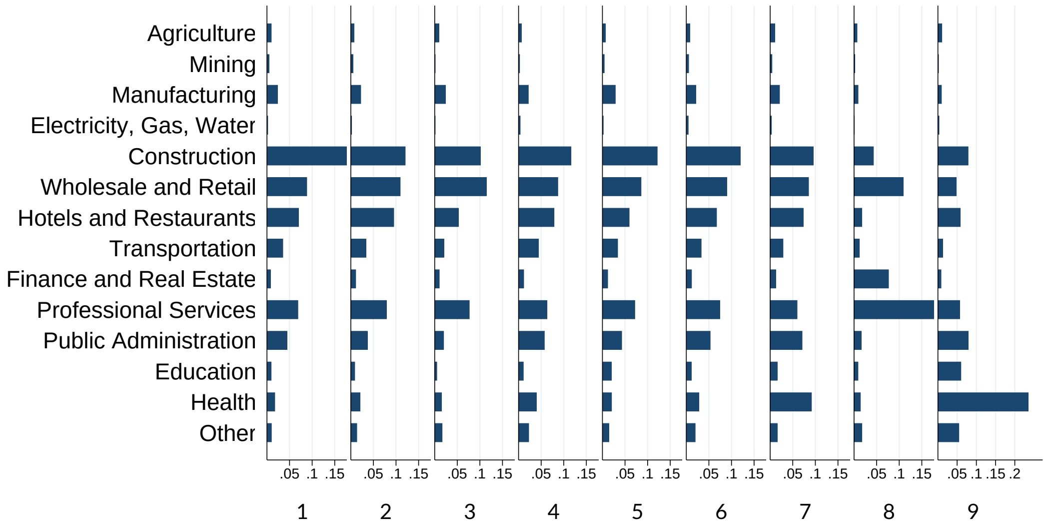


Figure 1.9: Histogram of Industry Composition by Market (1975-2005)

In Figure 1.10, I show an inverse relation between geographical and industry concentration. Concentration is measured using the popular Ellison-Glaeser concentration index (Ellison and Glaeser, 1997) which facilitates the comparison of regional and industry concentration between different markets.[24] In general, values for regional concentration are much higher than values for industry concentration. Moreover, high values of geographic concentration coincide with low values of industry concentration while markets that are more scattered around the country tend to be more specialized in specific industries.

[24] For a given market u, the Ellison-Glaeser index of concentration within R regions (or industries) is

$$EG_u = \frac{\sum_{r=1}^{R}(s_r - x_r)^2 - \left(1 - \sum_{r=1}^{R} x_r^2\right) H_u}{\left(1 - \sum_{r=1}^{R} x_r^2\right)\left(1 - H_u\right)},$$

where s_r denotes the share of market u employment in region (or industry) r, x_r denotes the share of total employment in region (industry) r and H_u is the Herfindahl index of the market firm size distribution, i.e., $H_u = \sum_{i:z_i=u} \left(\frac{\text{employment in } i}{\text{employment in } u}\right)^2$.

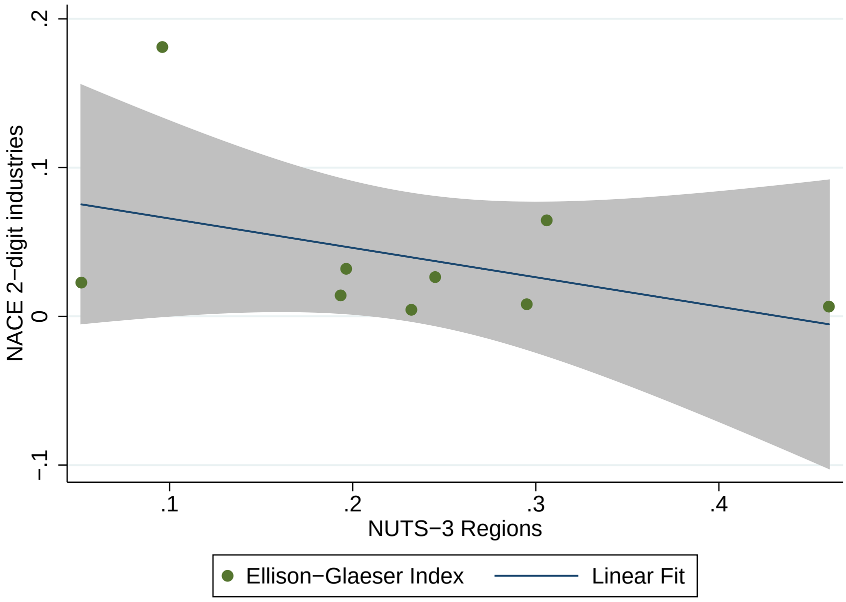


Figure 1.10: Ellison-Glaser-Index Regional and Industry Concentration (1975-2005)

Time Trends in the Geography of Endogenous Labor Markets over 1975-2005

In this section, I track developments in the structure of endogenous labor markets over time by estimating the SBM based on job mobility networks from the shorter sampling periods 1975-1980 until 2000-2005. This flexibility further distinguishes endogenous labor markets from the fixed nature of predefined local markets.

Figures 1.A1 to 1.A6 in the appendix display maps of the regional structure for job mobility networks for the (overlapping) six-year periods from 1975-1980 to 2000-2005. In general, they show a striking persistence in regional characteristics of the endogenous labor markets. Moreover, with exception of the early period from 1975-1980, there is a clear trend of increasing geographical mobility as labor markets are more and more scattered across several regions.

Table 1.4: Firm Distances within Labor Markets in km

	mean	sd.	1st quartile	median	3rd quartile
1975-1980	90.38	98.49	13.91	56.35	137.58
1980-1985	79.26	91.50	14.25	54.73	113.14
1985-1990	81.24	92.38	14.25	54.73	117.66
1990-1995	90.80	103.31	24.99	55.08	124.47
1995-2000	94.08	94.92	27.16	66.34	144.34
2000-2005	103.14	97.83	30.42	73.72	153.34

Note: Distance between firms is calculated according to the geographical distance between the centroid of the respective political districts.

This is also supported by an increase in the average distance between firms within labor markets. Table 1.4 shows aggregate statistics for the distribution of distances between all pairs of firms in the same labor market.[25] The average distance between firms within labor markets is 90.4 km in the late 1970s and 103.2 km in the early 2000s. After an initial decrease, it increased by 30% from the 1980s to the 2000s. The median distance increased by 31% from 56.3 km to 73.7 km over time. The largest increase can be found in the lower part of the distribution as the distance at the 1st quartile increased by 119%.

While geographical concentration is decreasing over time, the concentration of industries within endogenous labor markets increases slightly. The development of industry composition within endogenous markets over time is depicted in Figures 1.A7 to 1.A12 in the appendix. Moreover, Figure 1.11 compares the average Ellison-Glaeser index for geographical and industry concentration over all markets for each of the shorter sampling periods. After an initial increase between 1975-80 and 1980-85, geographical concentration steadily decreases over time while industry concentration exhibits an inverse pattern. Additionally, the error bars in Figure 1.11 indicate that the difference between both ways of concentration is statistically significant in earlier periods but becomes insignificant later on. These developments lend support to the hypothesis that over time individuals become more mobile (consistent with the larger size of labor markets in Table 1.4) and more specialized in specific industries.

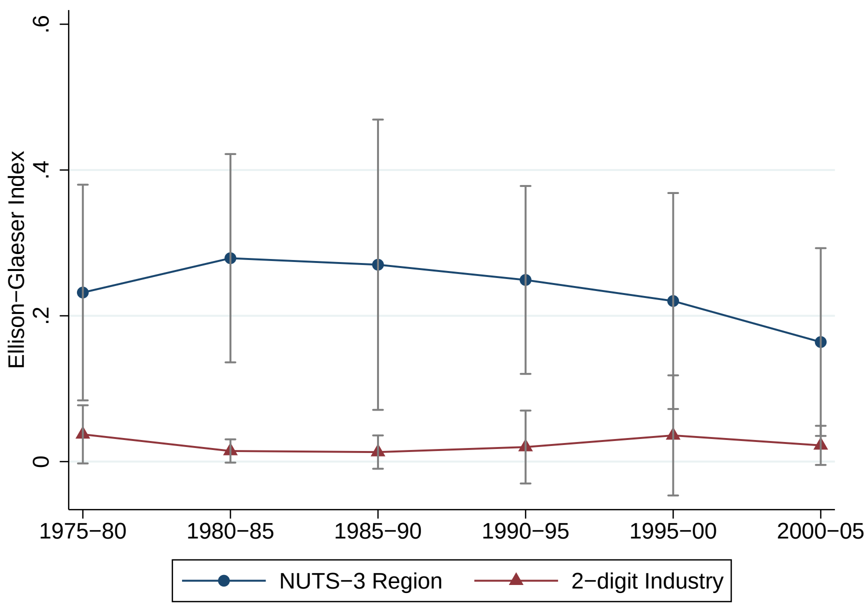


Figure 1.11: Ellison-Glaser-Index of NUTS-3 Region and Industry Concentration over Time

Summarizing the evidence so far, geographical factors seem to be the most important deter-

[25]Distances are computed by a relatively rough calculation where I assign to each firm the geographical coordinates of the centroid of the political district it is residing in. There are 95 political districts in Austria. The distance between firms in the same political district is underestimated as it is set to zero. The distance between firms in different political districts can be both under- or overestimated depending on the relative location to the centroid.

minant for the emergence of distinct labor markets. Additionally, largely separated markets in the same region differ by their industry or wage structure. Importantly, however, there is still substantial overlap in the distribution of observed characteristics such as regions, industries, and wages between markets. The endogenous labor market measure proposed in this paper allows to capture these unobserved determinants that drive frequent worker transitions between observationally different firms.

1.5.2 Worker Heterogeneity

In this section, I compare the scope of endogenous labor markets for various subgroups of the working force. The size and shape of labor markets can differ between worker types as the local availability of specific jobs varies and preferences towards mobility across regions and industries are heterogeneous. In the empirical analysis, I compare the outcome of the SBM estimated from job-to-job transitions separately by gender, nationality, age groups, and several skill measures.[26]

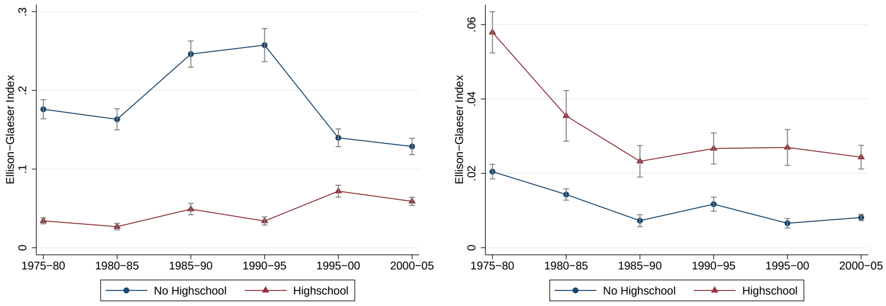
The most striking differences occur in the comparison of high-skilled and low-skilled workers. Figure 1.12 illustrates the development over time of endogenous labor markets defined by various measures of skill level. The graphs on the left refer to geographical concentration on the NUTS-3 level while the graphs on the right refer to 2-digit industry concentration, both measured by the average Ellison-Glaeser index over all markets. Remarkably, the different measures all point to the same conclusion: labor markets for higher skilled individuals are more dispersed in terms of geography but more specialized in specific industries than labor markets for low-skilled individuals. Panel 1.12a shows this difference for individuals without (blue dots) and with a highschool degree (red triangles), Panel 1.12b confirms that the same is true when considering individuals below (blue dots) and above the median (red triangles) in the distribution of individual fixed effects from an AKM wage decomposition, and Panel 1.12c shows the same pattern for blue collar (blue dots) versus white collar workers (red triangles). Difference between skill groups in the geographical scope of endogenous labor markets are also expressed in the geographical distances between firms in the same labor markets. Table 1.5 displays aggregate statistics of the distribution of distances (in km) between firms. On average, endogenous markets for white collar workers are about 25% bigger than endogenous markets for blue collar workers. The difference is even stronger between different schooling degrees. The average distance between firms in the same endogenous market is 69 km for individuals without highschool degree, 115 km for individuals with highschool degree, and 135 km for individuals with a university degree.

Finally, there is a similar difference for workers whose transition is associated with a wage

[26]Note that the modularity maximizing number of markets could be estimated differently in the subgroups making a comparison of the market structure more difficult. I therefore fix k to 9, the number that maximizes modularity in the full job mobility network.

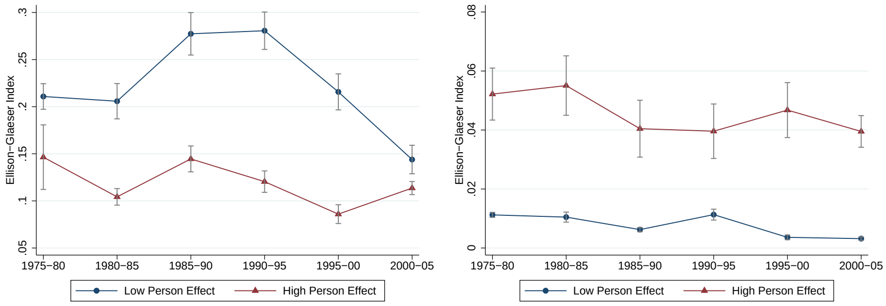
Figure 1.12: Concentration Indices for Markets based on Job-to-job Transitions of Subgroups

(a) Education



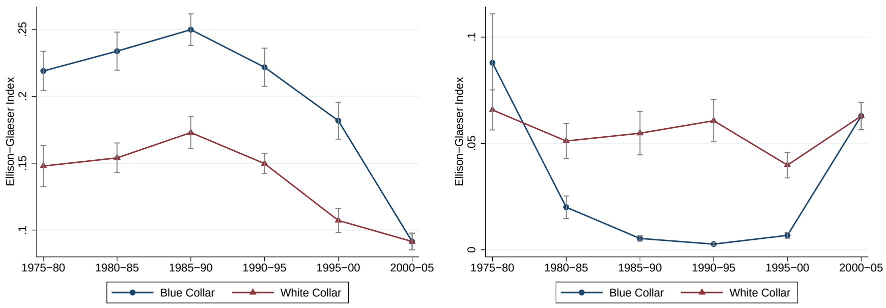
NUTS-3 2-digit Industry

(b) Individual Fixed Effect



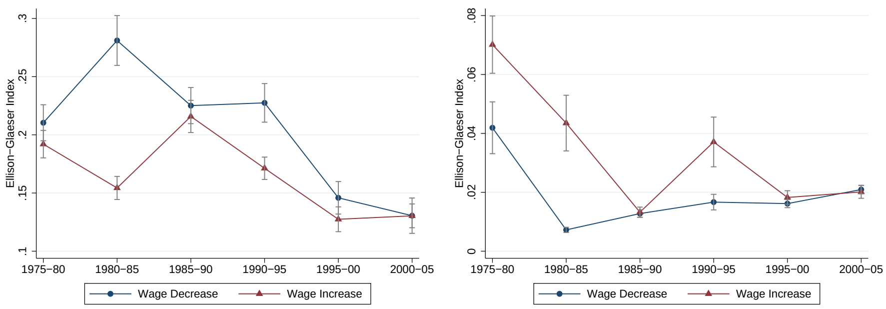
NUTS-3 2-digit Industry

(c) Occupation



NUTS-3 2-digit Industry

(d) Wage Change



NUTS-3 2-digit Industry

Table 1.5: Firm Distances within Labor Markets in km (1975-2005)

	mean	sd.	1st quartile	median	3rd quartile
Occupation					
blue collar workers	77.26	93.97	14.25	48.84	101.10
white collar workers	96.77	100.77	26.80	67.23	144.34
Education					
no highschool degree	68.90	81.08	0	45.56	96.58
highschool degree	115.12	111.25	16.69	100.95	178.13
university degree	135.21	121.77	16.69	141.60	201.98
Worker fixed effect					
below median	72.41	80.41	16.69	53.25	98.91
above median	94.86	101.43	26.30	62.46	140.97

Note: Distance between firms is calculated according to the geographical distance between the centroid of the respective political districts.

increase compared to those who incur a loss (Panel 1.12d). In particular, wage increases are associated with low regional concentration but higher industry concentration. The pattern is reversed for workers who incur wage cuts through the transition. Here regional concentration is higher than industry concentration. A potential interpretation pertains to job specialization which might be rewarded with high premiums while regionally less flexible workers incur wage cuts.

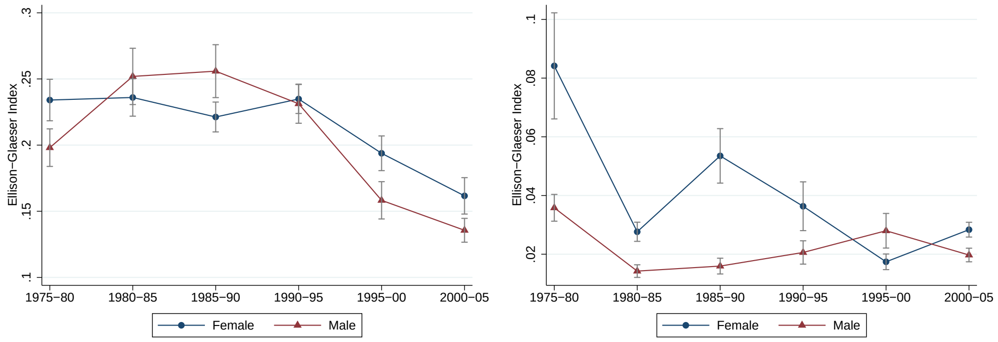
In figure 1.13, I display the development over time of endogenous labor markets defined by gender, nationality, and age. Again, the graphs on the left refer to geographical concentration on the NUTS-3 level while the graphs on the right refer to 2-digit industry concentration, both measured by the average Ellison-Glaeser index over all markets. Consistent with the aggregate trends described in section 1.5.1, there is a clear decrease in geographical concentration for all subgroups while industry concentration exhibits no strong direction.

Panel 1.13a indicates that there are no clear gender differences in the geographical concentration of labor markets. In terms of industry affiliation, however, labor markets of women (blue dots) are more specialized into specific industries than markets of men (red triangles) in the early periods until they converge in the late 1990s and 2000s. The situation is similar for Austrian (red triangles) versus Non-Austrian (blue dots) workers. Panel 1.13b shows no significant difference between geographical concentration in both subgroups. In early years, labor markets defined by Austrians were more concentrated within specific industries with convergence in later years.

Panel 1.13c compares endogenous labor markets for three age groups: young workers below 30 years of age (blue dots), middle-aged workers from 30 to 50 (red triangles), and elder workers above 50 (green squares). From the late 1980s on, middle aged workers are the most

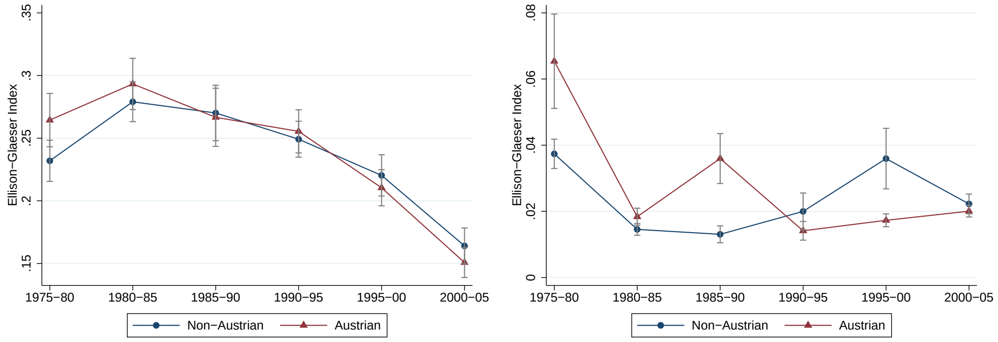
Figure 1.13: Concentration Indices for Markets based on Job-to-job Transitions of Subgroups

(a) Gender



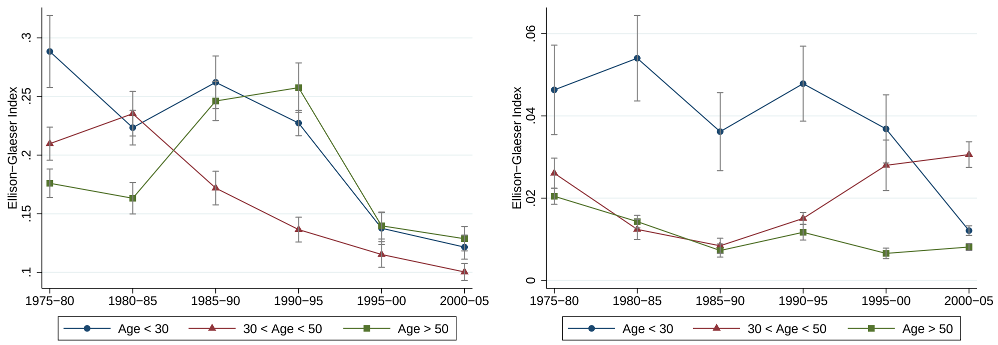
NUTS-3 2-digit Industry

(b) Nationality



NUTS-3 2-digit Industry

(c) Age Groups



NUTS-3 2-digit Industry

geographically mobile age group with least concentrated markets. Geographical concentration is higher for both younger and older workers. In terms of industry concentration, young workers have the most specialized markets for most of the time. Markets of middle-aged and older workers have low levels of industry concentration with a slight increase for middle-aged workers in later periods.

1.6 Mobility Responses to Economic Shocks

Based on my novel method to endogenously determine labor markets, I analyze mobility responses to large economic shocks. In particular, I use the SBM to predict the reactions to both, local and global shocks that hit specific parts of the economy.

The analysis proceeds in two parts. In the first application, I examine spillover effects of a large local labor demand shock, the breakdown of the Austrian steel industry in the late 1980s. I document negative spillover effects on employment in other firms from the same endogenous labor market both within and outside of the geographical location of the shock. In contrast, employment in firms from other endogenous markets within the affected region is largely unaffected by the shock.

In the second application, I use endogenous labor markets to predict mobility responses to global trade shocks, particularly the rising import competition from China and Eastern Europe. Previous research has identified the importance of job mobility to mitigate the negative consequences of global trade shocks (Autor et al., 2014; Dauth et al., 2016). An important and policy-relevant question however is *where* workers go when they are hit by trade shocks. The analysis shows that endogenous labor markets, estimated in the period *prior* to the shock, can accurately predict mobility responses while markets based on geographical areas fail to explain substantial parts of these movements.

In both applications, I take advantage of the flexibility of the SBM and vary the resolution of the analysis by modifying the number of markets, k. This allows me to quantify the effects at different levels of aggregation.

1.6.1 Local Labor Demand Shocks

The breakdown of the Austrian steel industry at the end of the 1980s was a particularly large shock that hit the Austrian economy unexpectedly. After World War II, Austria had nationalized its iron, steel, and oil industry in a protectionist act fearing expropriation by the Russian Army. The steel sector was mainly organized in one large company, the VÖEST. Mismanagement led to serious financial problems already starting in the mid-1970s. For several years, however, the Austrian government covered these losses. In November 1985, a big oil speculation scandal as well as the failure of a gigantic US plant project lead to an immediate turnaround in the company's strategy. The government installed a new management and enacted a strict restructuring plan with big mass layoffs and plant closures.[27]

Employment Spillovers

Mass layoffs that affect large companies can lead to substantial spillover effects on the surrounding economy (Gathmann et al., 2016). On the one hand, adverse shocks to labor demand can trigger domino effects that cause a downturn in the *local* economy due to decreasing demand for local goods and services or due to a negative impact on agglomeration economies. On the other hand, spillovers can affect economically connected firms (e.g., through input-output relations or common worker pools) that are not necessarily located in the same geographical area.

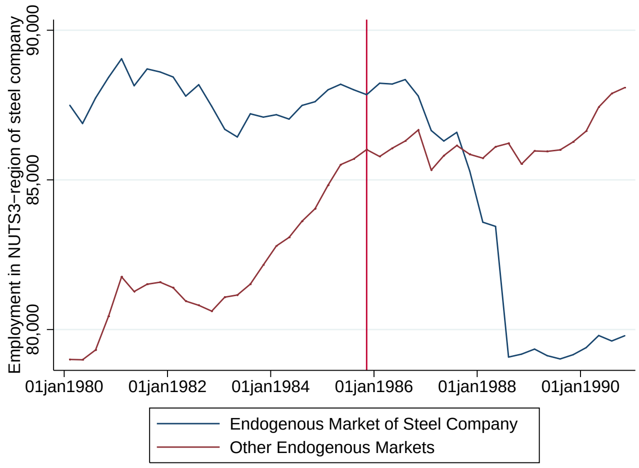
In order to examine which firms are negatively affected by spillovers through the breakdown of the VÖEST in Austria, I plot aggregate employment in all firms that are *not* affiliated to the steel sector over the time period from 1980 to 1990 in Figure 1.14. Panel 1.14a shows aggregate employment in the NUTS-3 region Linz-Wels where the VÖEST was located. I distinguish between employment in non-steel sector firms that are assigned to the same endogenous labor market as the steel company by the SBM (the blue line) and employment in non-steel sector firms from other endogenous markets (the red line).[28] In order to account for potential simultaneity in the determination of endogenous markets and responses to the shock, I estimate the SBM based on job mobility in the 5 years prior to the shock, 1980-1985.[29] The onset of the shock is indicated by the vertical line in November 1985. About one year after the strict restructuring plan in the VÖEST was enacted, employment in other (non-steel sector) firms from the same region and the same endogenous market sharply deteriorated. At the same time, the upward trend in employment at other firms of the same region but different endogenous markets continued almost at the same pace. Though I do not claim causality of these effects, the evidence suggests that – in line with the descriptive evidence in section 1.5 – there are several, independent markets within the same region. The connections detected and predicted by the SBM are capturing the boundaries that are relevant for the transmission of economic shocks.

This is also emphasized by the evidence in Figure 1.14b, where I plot aggregate employment in all (non-steel sector) firms from the endogenous market of the steel company that are lo-

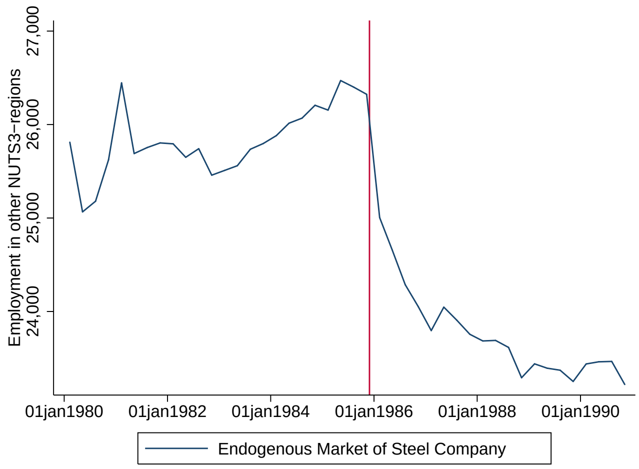
[27] Three years later, the trouble in the steel industry lead to an endogenous policy reaction, the massive extension of unemployment benefits in the Regional Extended Benefits Program from 1988. A series of paper is concerned with various effects of this policy change (Lalive and Zweimüller, 2004; Lalive, 2008; Lalive et al., 2015). Rather than looking at the effects of the policy response, the present application deals with the direct labor market effects of the breakdown.

[28] For comparison with the NUTS-3 regional classification, the SBM is estimated with $k = 35$ endogenous markets. The results also hold for other choices of k.

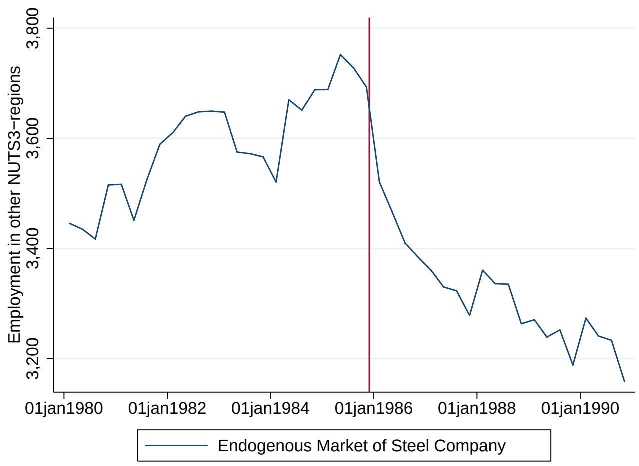
[29] Note that this implicitly assumes that assignments to endogenous labor markets are fixed in the short run. Moreover, new firms that are founded after 1985 are therefore excluded from the analysis.



(a) Employment in Non-steel Firms in the Region of the Steel Shock by Endogenous Labor Market



(b) Employment in Non-steel Firms from Other Regions but the Endogenous Labor Market of the Steel Shock



(c) Employment in Non-steel Firms from Non-REBP Regions but the Endogenous Labor Market of the Steel Shock

Figure 1.14: Spillovers on Employment in Non-steel Sector Firms

cated *outside* the NUTS-3 region Linz-Wels. The graph shows a strong decline in employment that starts exactly at the onset of the restructuring plan. Table 1.A2 in the appendix reports the 3-digit industry affiliations of these firms. They operate in economically related industries that rely on steel products (such as construction and manufacturing of motor vehicles) but are not directly involved in the same activities as the steel company. As a robustness check, I also plot in Figure 1.14c employment in (non-steel sector) firms from the same endogenous labor market that are located outside the REBP area defined in Lalive et al. (2015) which excludes other areas that focus on steel-related industries.[30] The pattern of rapid decline after the onset of the shock is unchanged. The evidence therefore suggests that the SBM can also detect and predict relevant economic ties to other firms outside the steel sector and the original location.

Combining the evidence from both panels, I find strong indications that spillovers to other firms operate on the level of economic ties rather than through local multipliers. This is in line with the evidence in Gathmann et al. (2016).[31] Endogenous labor markets estimated by the SBM seem to be better suited to detect these economic connections than fixed geographical boundaries.

Mobility Responses

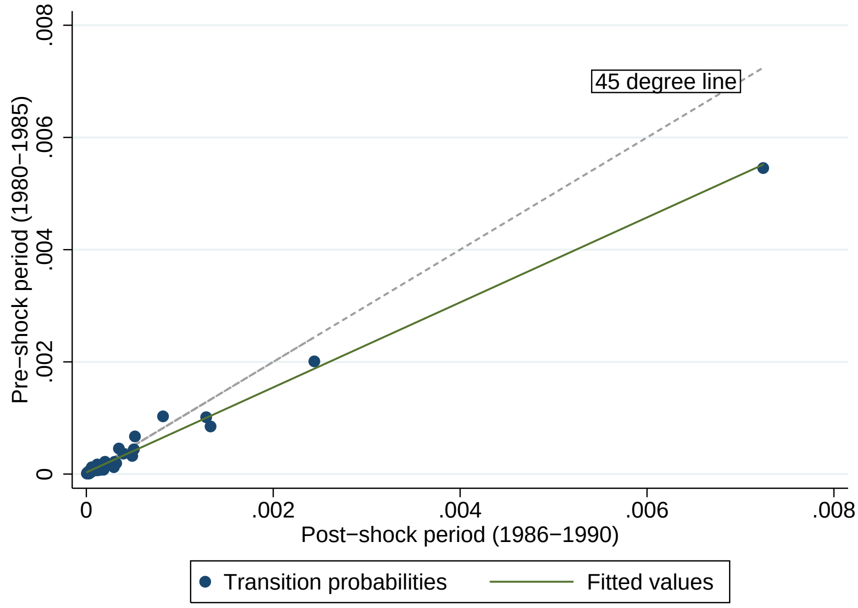
In this section, I briefly examine the effect of the local demand shock on the transition probabilities between the endogenous labor markets. On the worker level, the negative impact on employment in the endogenous labor market that contained the steel company can be partly offset by job mobility to other markets.

After the shock, I find an increase in the share of job-to-job transitions out of the endogenous labor market containing the steel company and a decrease in the share of job-to-job transitions into this market. The share of transitions away from the affected market to other endogenous markets among all transitions in the economy increases from 1.5 to 1.9 percent. At the same time, transitions into the affected market decrease from 1.5 to 1.2 percent of all job-to-job transitions.

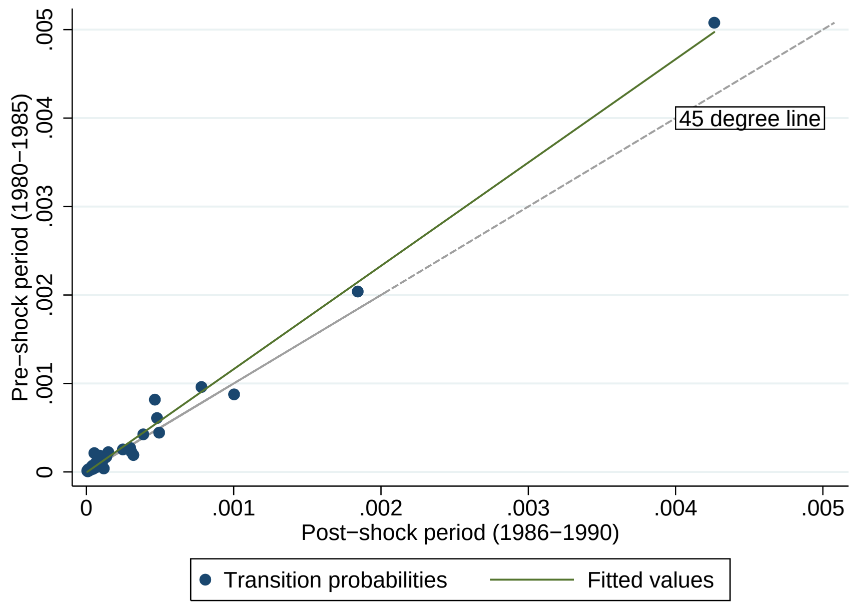
Most importantly, the impact on transition probabilities to other markets is larger the higher the initial transition probability was in the period before the shock. Figure 1.15 shows a scatter plot of transition probabilities in the period prior to the shock (1980-1985 on the vertical axes) plotted against transition probabilities after the shock (1986-1990 on the horizontal axes). Comparison to the dashed 45 degree line indicates that there is a change in the composition of transitions between these periods. In the upper panel, I illustrate the probability of transitions from the affected endogenous market into the other 34 markets. Workers who leave the affected market increasingly target those markets that had a stronger connec-

[30] Besides the main center in Linz, parts of the Austrian steel industry were located in various parts of Styria.

[31] Similarly to their paper, I do not find significant impacts of the labor demand shock on wages. This could potentially be explained by downward wage rigidity.



(a) Transitions out of Shocked Market



(b) Transitions into Shocked Market

Figure 1.15: Transition Probabilities Before and After the Labor Demand Shock in the Endogenous Market of the Steel Company

tion before the shock. The opposite pattern can be found for transitions into the affected endogenous market in panel (b). Transitions into this market become less likely, especially those from markets with a high pre-shock transition probability. The evidence suggests that an adverse shock to local labor demand leads to changes in job mobility that are roughly proportional to the transition probabilities predicted by the SBM.

In summary, endogenous labor markets estimated by the SBM help to identify the relevant parts of the economy that are affected by spillover effects of adverse economic shocks. Moreover, they help to predict job mobility flows in response to such shocks. In the following section, I further analyze the policy-relevant question of worker reallocation by examining a different type of shock, the increase in import competition from eastern countries.

1.6.2 Global Trade Shocks

The unprecedented rise in the importance of China and Eastern Europe for global trade over the past decades has caused strong disruptions in the job biographies of workers in industrialized countries (Autor et al., 2013, 2014; Dauth, Findeisen, and Suedekum, 2014; Dauth et al., 2016). Import competition through largely exogenous shifts in the productivity of eastern countries triggered a rapid decline in wages and employment for workers in affected manufacturing industries, both in the US and Germany.[32][33]

Workers who suffer from wage or employment losses through strong exposure to import competition can mitigate the negative impact by switching their employer, industry, sector, or region. In the present application, I am particularly interested in the relative importance of different margins of mobility. Specifically, I examine whether the SBM introduced in section 1.4 is able to predict the mobility responses following global trade shocks more accurately than ad-hoc definitions of local markets such as regional entities. To this aim, I augment the studies of Autor et al. (2014) and Dauth et al. (2016) by introducing endogenous labor markets estimated by the SBM. The analysis proceeds in two steps. First, following Autor et al. (2014) I decompose the causal impact of trade shocks on medium-run accumulated earnings into additive components that accrue within the original firm, region, industry, and endogenous labor market and through mobility between these units. Second, following Dauth et al. (2016) I estimate the contemporaneous impact of trade shocks on earnings using high-dimensional

[32] The fall of the iron curtain and the ensuing transition of eastern European countries into market economies can be considered a largely unexpected event. Similarly, the rapid improvements in China's competitiveness, also boosted by its entry into the WTO in 2001, are mainly driven by internal factors. Furthermore, I follow the common strategy in the literature and instrument Austria's exposure to trade with Eastern countries using trade exposure of other high-income countries in order to account for possible correlation between imports and domestic demand or productivity shocks. Detailed discussions of the identification strategy are provided by Autor et al. (2014) and Dauth et al. (2016).

[33] Export opportunities due to market liberalization in eastern countries, in contrast, increased wages and employment in specific industries in Germany while there seems to be no such offsetting effect for US manufacturing workers.

fixed effects to separate direct responses and mobility responses.

Both strategies show that endogenous labor markets estimated in the period *prior* to the shock are much better predictors of mobility adjustments after the shock than traditional concepts based on predefined characteristics. Workers with strong exposure to the shock mitigate the negative impact by switching to firms within the original labor market but in different industries and regions.

Data

Data on trade exposure I acquire data on trade exposure from the United Nations Commodity Trade Statistics Database (Comtrade) which provides annual import and export statistics of over 170 reporter countries detailed by commodities and trade partner countries. I obtain Austrian trade data on the SITC3 5-digit commodity level and merge it to the NACE95 3-digit industries in the ASSD using correspondence tables provided by the World Bank.[34]

Figure 1.16 demonstrates the growing importance of imports from the East compared to total imports.[35] Trade volumes are normalized to 1 in 1990 and shown on a log-scale. The solid red line indicates a tenfold increase in imports from the East between 1990 and 2010 for the median industry. In contrast, total imports from all countries have only doubled for the median industry as shown by the solid black line. The dashed lines illustrate the increase in imports for industries at the 25th and 75th percentile respectively and show that there is also more variation between industries in import exposure to the East.[36]

Exposure to imports from the East varies on the 3-digit industry level. For each worker i in industry $j(i)$, import exposure in year t is measured by

$$ImE_{j(i),t} = 100 \times \frac{IM^{EAST \rightarrow AUT}_{j(i),t}}{\sum_{\ell:j(\ell)=j(i)} w_{\ell,t-1}}, \tag{1.10}$$

where, $IM^{EAST \rightarrow AUT}_{j,t}$ denotes aggregate Austrian imports from the East in industry j and year t. Imports are normalized by the initial size of industry j in the Austrian economy, measured by the total wage bill in the previous year, $\sum_{\ell:j(\ell)=j(i)} w_{\ell,t-1}$.

[34] As in Dauth et al. (2016), ambivalent cases are partitioned according to Austrian employment shares in 1978. Moreover, I convert all trade values into 2010-Euros using historical exchange rates provided by the Austrian National Bank and the Austrian CPI.

[35] The countries subsumed in the East comprise Bulgaria, China, the Czech Republic, Hungary, Poland, Romania, Slovakia, Slovenia, Russia, Belarus, Estonia, Hongkong, Latvia, Lithuania, Macau, Moldova, Ukraine, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan.

[36] Figure 1.A13 in the appendix shows a similar picture for exports. The rise in Austrian exports to the East is much less pronounced than for imports. In the empirical analysis, I therefore focus on the exposure of the Austrian economy to imports from the East. Tables 1.A3 and 1.A4 in the appendix report the industries with the largest increase in im- and exports respectively.

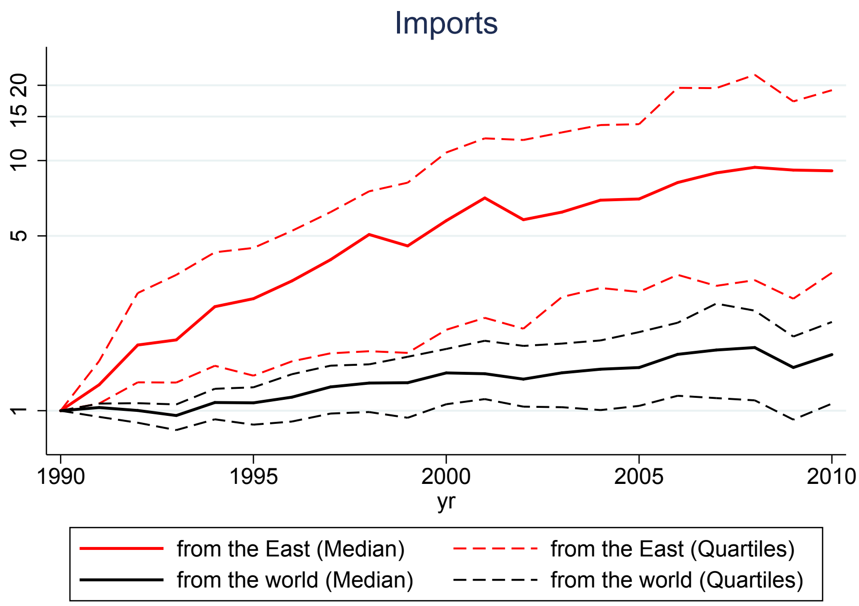


Figure 1.16: Rising Import Volumes in Austrian Trade

Panel of Manufacturing workers I merge the information on trade exposure to individual level data on workers from the ASSD in the time period from 1990 - 2010. Following Dauth et al. (2016), I split the data into two balanced ten-year panels with base years 1990 and 2000. Each panel consists of all individuals who are between 22 and 54 years old and have their main job (i.e., the job spell with the longest duration) in the manufacturing sector in the base year.[37] For each worker, I track the job biography over the 10 year period and compute the sum of annual earnings for each year (which could be zero due to non-employment in some years).[38] In case of multiple job spells within a year, regional, industry, and labor market information refers to the main job. In case of non-employment, these characteristics are taken from the last employment spell, assuming some short term attachment to regions, industries and markets. Most importantly, the assignment of firms to endogenous labor markets is estimated based on worker flows in the 5-year period prior to the shock (1985-1990 for the base year 1990 and 1995-2000 for the base year 2000) in order to account for potential simultaneity in the formation of the network and the mobility responses to trade shocks.

Table 1.6 presents summary statistics for earnings and trade exposure of workers separately for both periods. The first row in each panel characterizes the distribution of base year earnings in the sample. The second row shows accumulated earnings for the ten-year period relative to the base year earnings level. The median worker received exactly 10 times his base-year

[37] Individuals who die within 10-year period are dropped.

[38] Since wage information in the ASSD is censored at the social security contribution limit, I merge the data to uncensored tax records from the Austrian Ministry of Finance. Uncensored information, however, is only available since 1995. For earlier periods, I therefore impute the upper tail of the wage distribution using a strategy similar to the one used in Card et al. (2013).

Table 1.6: Descriptive Statistics

	mean	sd.	1st quartile	median	3rd quartile
Panel A. 1990-2000					
Earnings					
base year	29798.14	18324.82	19351.46	27489.76	36986.72
accumulated / base year	1182.09	4826.94	611.66	1003.99	1188.88
yearly / base year	116.55	501.75	68.76	100.81	117.84
Δ Import Exposure					
in base year industry	.257	.474	.031	.143	.329
yearly	.012	.193	-.003	.003	.028
Observations	499,706				
Panel B. 2000-2010					
Earnings					
base year	34823.11	19969.02	23452.34	31808.98	42130.67
accumulated / base year	1196.39	5307.30	811.53	1026.18	1167.24
yearly / base year	117.85	540.70	89.54	101.15	115.33
Δ Import Exposure					
in base year industry	.325	.835	.078	.186	.295
yearly	.008	.346	-.009	.004	.049
Observations	436,735				

Note: The change in import exposure is measured on the 3-digit industry level and computed by the change in imports from the East normalized by the (lagged) wage bill. Results are derived from 5,496,766 yearly observations of 499,706 workers in Panel a. and 4,804,085 observations from 436,735 workers in Panel B. workers. Base year earnings are expressed in 2010 Euros. Accumulated Earnings are added over the entire period and normalized by base year earnings. The change in import exposure is computed for the base year industry over the entire period and on a yearly base for the current industry.

earnings for the period from 1990 to 2000 while there is a substantial degree of (right-skewed) variation around the median. Values for the 2000s are slightly higher and more compressed. The third row characterizes the distribution of yearly earnings relative to the base-year level. The median amounts to 100% of base year earnings, again with substantial variation. In particular, the first quartile of yearly relative earnings is only 69% (89% in the second period) of the base level while the third quartile of yearly earnings amounts to 117% (115%) of base-year earnings. There is also substantial variation in the individual exposure to imports from the East. The median change in the eleven-year difference of equation (1.10) is 0.143 (0.186) while the difference in exposure is 0.031 (0.078) for workers at the first quartile and 0.329 (0.295) for workers at the third quartile. Similarly, the yearly change in the exposure measure in row 5 shows substantial variation across workers.

Medium-run Analysis

In the first estimation strategy, I follow Autor et al. (2014) and estimate the impact of the eleven-year difference of trade exposure in the base year industry on accumulated earnings over the entire period (relative to base year earnings). I pool both panels and estimate the following model:

$$Y_{i\tau} = \beta_0 + \beta_1 \Delta ImE_{j(i),\tau} + \mathrm{x}'_{i\tau}\alpha + \varphi_{J(i),\tau} + \varphi_{R(i),\tau} + \varphi_{M(i),\tau} + \varphi_\tau + \varepsilon_{i\tau}, \tag{1.11}$$

where $Y_{i\tau} = \sum_{t=\tau+1}^{t=\tau+10} \frac{Y_{it}}{Y_{i\tau}}$ denotes accumulated earnings relative to base year earnings for $\tau \in \{1990, 2000\}$ and $\Delta ImE_{j(i),\tau} = ImE_{j(i),\tau+10} - ImE_{j(i),\tau}$ denotes the change in import exposure in the industry where i was employed in base year τ. Additional controls subsumed in x_i, are indicators for female gender and foreign born status, for 7 different age categories, 3 different occupation categories, 3 different tenure groups, and for 5 different groups of firm size in the base year.

Identification of causal effects in this model is extensively discussed in Autor et al. (2014). I follow their strategy and instrument import and export exposure using trade flows between other countries and the East in order to purge the effect of domestic shocks within Austria that simultaneously affect trade and labor market outcomes.[39] Moreover, the model includes dummies for broad manufacturing industries, $\varphi_{J(i),\tau}$, states, $\varphi_{R(i),\tau}$, and endogenous labor

[39] In particular, import exposure is instrumented by trade flows of other (non-neighboring) developed countries which are not in the Euro zone,

$$ImE^{INSTR}_{j(i),t} = \frac{IM^{EAST\rightarrow INSTR}_{j(i),t}}{\sum_{\ell:j(\ell)=j(i)} w_{\ell,t-3}}$$

where *INSTR* comprises Australia, New Zealand, Japan, Singapore, Canada, Denmark, Sweden, Norway, and the UK. Note that the normalization now contains the wage bill in $t-3$ in order to account for sorting across industries in anticipation of future trade flows with China.

markets, $\varphi_{M(i),\tau}$, in order to control for potentially different industry-, state-, or market-level trends. Finally, the dummy φ_τ separates the two ten-year panels.

The main estimate for the model in equation (1.11) is shown in the first column of Table 1.7. There is a strong and (weakly) significant negative impact of the eleven-year change in import exposure on accumulated earnings.

Table 1.7: Estimation Results - Accumulated Earnings and Job Switch Probabilities

	(1)	(2)	(3)	(4)	(5)	(6)
	all employers	initial firm	endogenous labor market			
			same		other	
Panel A. Accumulated Wages						
same 3-digit industry			yes	no	yes	no
ΔImE	-48.11**	-59.90***	-31.29	48.07***	-8.820***	3.835
	(23.04)	(21.73)	(19.55)	(18.06)	(2.333)	(16.91)
same NUTS-3 region			yes	no	yes	no
ΔImE			5.707	11.06	-5.984	0.998
			(20.61)	(10.11)	(14.52)	(7.521)
Panel B. Job Switch Indicator						
same 3-digit industry			yes	no	yes	no
ΔImE			-0.045***	0.027***	-0.006***	0.024***
			(0.013)	(0.009)	(0.002)	(0.009)
same NUTS-3 region			yes	no	yes	no
ΔImE			-0.035***	0.018**	0.008	0.010**
			(0.011)	(0.008)	(0.005)	(0.005)
1st stage F	12.166					
N	936,392					

Note: Clustered standard errors on the industry times base year level in parentheses. Results are reported for 2SLS estimates of equation (1.11) where the change in import exposure is instrumented with the corresponding change of exposure in other high-income countries. The decomposition of the total effect is additive such that the difference between the aggregate effect in column (1) and the effect in the initial firm in column (2)results from the sum of the effects in columns (3) to (6). Endogenous labor markets are estimated in the 5 years before the base year with $k = 35$ markets.

The main purpose of the analysis is to decompose this total effect of trade exposure on accumulated earnings into additive parts that capture the direct effect of the shock (excluding mobility responses) as well as the different mobility margins. Column 2 of Table 1.7 shows estimates for the effect of trade exposure on all earnings that accrued in the initial firm which employed the worker in the base year. This effect is even more negative than the total effect, indicating that workers incur huge earnings and job losses in firms that are negatively affected by import exposure over the eleven-year period. They can however partly make up for these losses by switching to different firms, industries, regions, or labor markets. Columns (3) to (6) display different types of mobility responses within and between 3-digit industries, NUTS-3 regions, and endogenous labor markets with $k = 35$. The estimates referring to industry and market mobility in the first row indicate that workers with larger shocks generate significantly less earnings from firms in the same 3-digit industry (columns 3 and 5). They have however significantly higher earnings from firms in the same endogenous market but different industries (column 4). The impact on earnings from outside the original labor market and industry is small and insignificant. The estimates referring to regional and market mobility in the second row are very noisy. The direction of the effects however suggests that individuals with a stronger exposure to the shock generate more earnings from other firms in the same endogenous labor market (columns 3 and 4) but less earnings from firms in the same region but in a different labor market (column 5). Panel B of Table 1.7 displays the results for a related analysis where I replace the outcome variable with job switch indicators and estimate linear probability models. Column (3) indicates that workers with stronger import exposure have a lower probability to stay in their original 3-digit industry and endogenous labor market. There is however a positive and significant effect on the probability to switch the industry but to remain in the original endogenous labor market (column 4). Conversely, the probability to switch the labor market but to remain in the original industry is slightly negatively affected by the shock. The probability to switch both, endogenous market and industry, is also increased for highly exposed workers. The picture looks very similar for job switch probabilities between NUTS-3 regions and endogenous labor markets. In summary, mobility *between* regions and industries but *within* endogenous labor markets appears to be a main mechanism to mitigate the negative impacts of exposure to import competition from eastern countries.

Short-run Analysis

In the second empirical strategy, I address the caveat that in the medium-run model in equation (1.11) all outcomes are related to the trade shock in the initial industry. As mobility responses are an important aspect of adjustments to trade shocks, it is important to examine the impact of the actual contemporaneous exposure to trade in the current industry on earnings. To this aim, I follow Dauth et al. (2016) in estimating an annual panel model,[40]

$$Y_{it} = \beta_0 + \beta_1 \cdot ImE_{j(i),t} + x_{it}'\alpha + \varphi_{t,J(j)} + \varphi_{t,R(i)} + \varphi_{t,M(i)} + \gamma_i + \varepsilon_{it}. \tag{1.12}$$

The most important difference to the medium-term model in equation (1.11) is the inclusion of individual level fixed effects, γ_i. Hence, the effect of trade exposure on earnings is identified on variation *within* individuals rather than *between* individuals with common observable characteristics. Estimation results for the baseline version of equation (1.12) are displayed in column 1 of Table 1.8. The total effect of contemporaneous import exposure on annual earnings is significantly negative confirming the medium-run evidence in the previous analysis. It

[40] Again, import exposure is instrumented with the relevant exposure from other countries.

is identified by within variation in earnings due to wage changes and non-employment as well as by variation due to job mobility.

Table 1.8: Estimation Results – Short-run Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ImE	-0.548***	-1.140***	-1.071***	-0.567***	-0.682***	-0.544***	-0.560***
	(0.111)	(0.205)	(0.196)	(0.112)	(0.124)	(0.107)	(0.118)
FE	i	$i \times firm$	$i \times ind.$	$i \times state$	$i \times NUTS3$	$i \times market$ $(k = 9)$	$i \times market$ $(k = 35)$
R^2	0.782	0.910	0.887	0.796	0.841	0.800	0.857
Groups	936,441	1,532,792	1,346,855	1,029,934	1,161,066	984,332	1,101,996
KP	656.1	40.57	39.24	554.9	472.7	580.8	438.5

Note: 10,300,851 observations of 936,441 workers. The main regressor is import exposure, ImE. Further controls include age polynomials, 1-digit-industry × year, state × year, and endogenous market × year dummies. Standard errors, clustered by industry × year in parentheses. KP denotes the Kleinberg-Papp statistic of the first stage. $^{***}p < 0.1$ $^{**}p < 0.05$, $^{*}p < 0.01$

In order to assess the relative importance of various margins of mobility responses to the contemporaneous trade shock, I replace the individual fixed effects by different sets of higher-dimensional fixed effects in equation (1.12). Particularly, including individual times firm-level fixed effects captures the *direct* effect of import exposure on earnings by exploiting only variation within spells in the same firm. The estimates in column (2) of Table 1.8 indicate that the negative impact is much stronger in this specification (consistent with the evidence from the medium-term analysis). The difference between the two estimates in columns (1) and (2) derives from mobility responses as the direct effect excludes variation that derives from firm switches. To examine which type of mobility responses helps to mitigate the adverse direct impact, I include fixed effects on the individual times industry level (column 3), individual times state level (column 4), individual times NUTS3-region level (column 5), and the individual times endogenous labor market level (column 6 with $k = 9$ labor markets and column 7 with $k = 35$ labor markets). Absorbing variation between industries into the individual times (3-digit) industry fixed effect (column 3) leads to a strongly negative effect that is similar to the direct effect in column 2. Movements between 3-digit industries are therefore very important to mitigate negative trade impacts. Columns 4 and 5 show that the estimates absorbing movements across states and NUTS3 regions are in between the direct and the aggregate effect. Mobility adjustments between NUTS3 regions are still an important part of wage responses while absorbing variation between states has almost no impact on the trade effect.

Similarly to the evidence from the medium-run analysis, it is variation on the endogenous labor market level that returns an the estimate closest to the aggregate effect. The estimates using only variation within the endogenous labor markets with $k = 9$ and $k = 35$ are very close to the aggregate effect. Moreover, Figure 1.17 shows that this is the case even for a finer disaggregations of endogenous labor markets. This confirms the striking ability of the SBM to predict labor market flows that mitigate the negative impact of global trade shocks. Even for a detailed view on the economy with 1000 labor markets (about 95 firms on average per market) the SBM accurately predicts those sets of firms that offer better employment and earnings possibilities to workers.

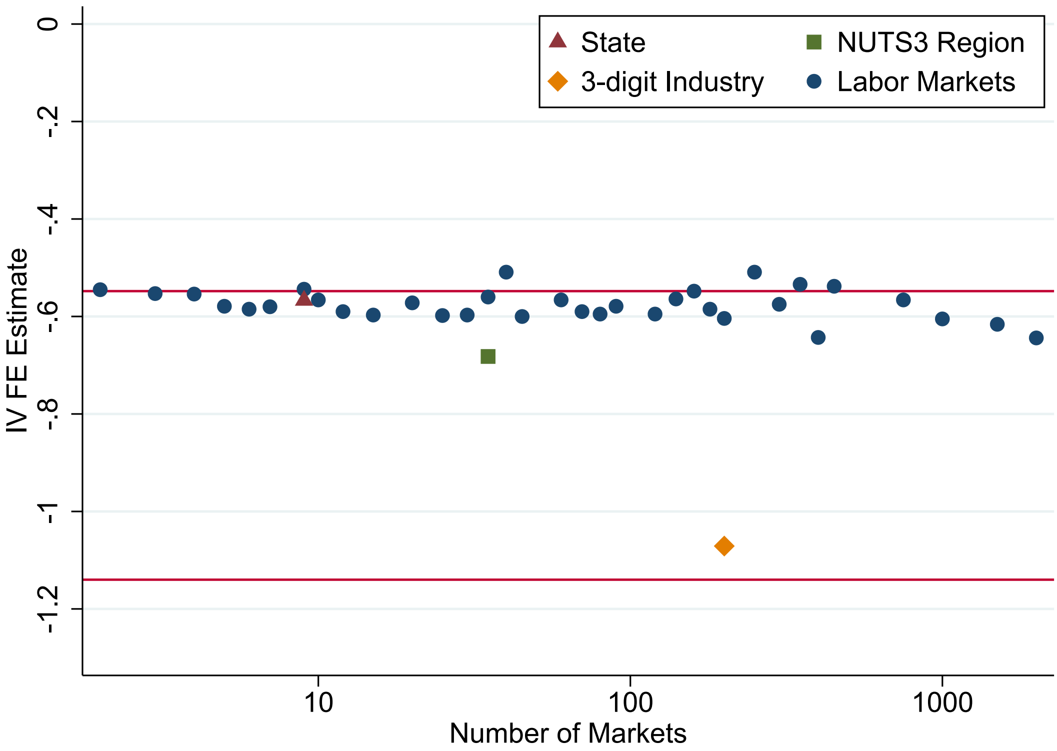


Figure 1.17: Impact of Trade Shock on Annual Earnings - Various Models

1.7 Conclusion

The Stochastic Block Model is an interesting novel device to enrich the toolbox of economists who work with network data. In the present context, it allows recovering endogenous labor markets in Austria from observed worker flows. These endogenous labor markets are geographically clustered but differ substantially from labor markets based on administrative borders. Furthermore, reflecting differences in mobility patterns, markets become more geographically dispersed over time and vary substantially across worker types.

The empirical analysis of job mobility responses to labor demand and trade shocks highlighted how endogenously determined labor markets can be used to better predict and understand worker flows in the economy. The increasing availability of administrative matched employer-employee data covering full populations should allow to apply my method to other countries and contexts. Interesting extensions such as migration induced labor supply shocks could therefore be addressed in future research. An important question for the future of the European Union regards the degree of between-country job mobility to balance inequalities. Endogenous international labor markets could provide answers to important policy questions in this context.

The SBM can be used to identify endogenous markets based on different kinds of networks. In the trade literature, for instance, researchers have analyzed the role of production networks (based on supplier relationships between firms) for aggregate outcomes (e.g., Carvalho, 2014; Chaney, 2014). Understanding the market structure in these networks might help to examine spillover effects between firms or countries that are not directly linked but exposed to similar market-level shocks.

Finally, several model extensions of the SBM have been introduced in the recent network literature. Airoldi, Blei, Fienberg, and Xing (2009) and Aicher, Jacobs, and Clauset (2015) consider mixed-membership models where nodes can belong to different communities depending on the kind of interaction. Peixoto (2014b) describes a hierarchical SBM where communities are nested in multiple levels. Finding consistent estimation methods and applying these models to economic networks is an interesting avenue for future research.

References

John M. Abowd, Francis Kramarz, and David N. Margolis. High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333, 1999.

John M. Abowd, Robert Creecy, and Francis Kramarz. Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Unpublished working paper, Cornell University Department of Economics, March 2002.

Daron Acemoglu and Joshua Angrist. How Large are Human-capital Externalities? Evidence from Compulsory-schooling Laws. In *NBER Macroeconomics Annual 2000, Volume 15*, pages 9–74. MIT Press, 2000.

Christopher Aicher, Abigail Z. Jacobs, and Aaron Clauset. Learning Latent Block Structure in Weighted Networks. *Journal of Complex Networks*, 3(2):221–248, 2015.

Edoardo M Airoldi, David M Blei, Stephen E Fienberg, and Eric P Xing. Mixed Membership Stochastic Blockmodels. In *Advances in Neural Information Processing Systems*, pages 33–40, 2009.

David H. Autor, David Dorn, and Gordon H. Hanson. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6):2121–2168, 2013.

David H. Autor, David Dorn, Gordon H. Hanson, and Jae Song. Trade Adjustment: Worker Level Evidence. *Quarterly Journal of Economics*, 129(4), 2014.

Ragnhild Balsvik. Is Labor Mobility a Channel for Spillovers from Multinationals? Evidence from Norwegian Manufacturing. *The Review of Economics and Statistics*, 93(1):285–297, 2011.

Peter J Bickel and Aiyou Chen. A Nonparametric View of Network Models and Newman–Girvan and other Modularities. *Proceedings of the National Academy of Sciences*, 106(50): 21068–21073, 2009.

Melissa Bjelland, Bruce Fallick, John Haltiwanger, and Erika McEntarfer. Employer-to-Employer Flows in the United States: Estimates Using Linked Employer-Employee Data. *Journal of Business & Economic Statistics*, 29(4):493–505, 2011.

Olivier Blanchard and Larry Katz. Regional Evolutions. *Brookings Papers on Economic Activity*, 1:1–75, 1992.

Stéphane Bonhomme and Elena Manresa. Grouped Patterns of Heterogeneity in Panel Data. *Econometrica*, 83(3):1147–1184, 2015.

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. A Distributional Framework for Matched Employer Employee Data. Unpublished manuscript, University of Chicago, 2016.

John Bound and Harry Holzer. Demand Shifts, Population Adjustments and Labor Market Outcomes during the 1980s. *Journal of Labor Economics*, 18:20–54, 2000.

Matias Busso, Jesse Gregory, and Patrick Kline. Assessing the Incidence and Efficiency of a Prominent Place Based Policy. *American Economic Review*, 103(2):897–947, 2013.

David Card, Jörg Heining, and Patrick Kline. Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128, August 2013.

Vasco M. Carvalho. From Micro to Macro via Production Networks. *Journal of Economic Perspectives*, 28(4):23–48, 2014.

Giacinta Cestone, Chiara Fumagalli, Francis Kramarz, and Giovanni Pica. Insurance Between Firms: The Role of Internal Labor Markets. Technical Report DP11336, CEPR, June 2016.

Thomas Chaney. The Network Structure of International Trade. *American Economic Review*, 104:3600–3634, 2014.

Aaron Clauset, Mark EJ Newman, and Cristopher Moore. Finding Community Structure in Very Large Networks. *Physical review E*, 70(6):066111, 2004.

Guido Matias Cortes and Giovanni Gallipoli. The Costs of Occupational Mobility: An Aggregate Analysis. *Journal of the European Economic Association*, forthcoming.

Bruno Crepon, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment. *Quarterly Journal of Economics*, 128(2):531–580, 2013.

Leon Danon, Albert Diaz-Guilera, Jordi Duch, and Alex Arenas. Comparing Community Structure Identification. *Journal of Statistical Mechanics: Theory and Experiment*, 2005 (09):P09008, 2005.

Kunal Dasgupta. Learning and Knowledge Diffusion in a Global Economy. *Journal of International Economics*, 87(2):323–336, 2012.

Wolfgang Dauth, Sebastian Findeisen, and Jens Suedekum. The Rise of the East and the Far East: German Labor Markets and Trade Integration. *Journal of the European Economic Association*, 12(6):1643–1675, 2014.

Wolfgang Dauth, Sebastian Findeisen, and Jens Suedekum. Adjusting to Globalization - Evidence from Worker-Establishment Matches in Germany. Technical report, University of Mannheim, 2016.

Christian Dustmann, Uta Schönberg, and Jan Stuhler. Labor Supply Shocks and the Dynamics of Local Wages and Employment. *Quarterly Journal of Economics*, forthcoming, 2016.

Andreas Dzemski. An Empirical Model of Dyadic Link Formation in a Network with Unobserved Heterogeneity. Technical report, University of Mannheim, 2014.

Jonathan Eaton and Samuel Kortum. Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779, 2002.

Glenn Ellison and Edward L. Glaeser. Geographic Concentration in US Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105:889–927, 1997.

Marc Ferracci, Grégory Jolivet, and Gerard J. van den Berg. Evidence of Treatment Spillovers Within Markets. *Review of Economics and Statistics*, 96(5):812–823, 2014.

Martina Fink, Esther Kalkbrenner, Andrea Weber, and Christine Zulehner. Extracting Firm Information from Administrative Records: The ASSD Firm Panel. Technical report, NRN Working Paper: The Austrian Center for Labor Economics and the Analysis of the Welfare State, 2010.

Santo Fortunato. Community Detection in Graphs. *Physics Reports*, 486(3):75–174, 2010.

Christina Gathmann, Ines Helm, and Uta Schönberg. Spillover Effects in Local Labor Markets: Evidence from Mass Layoffs. Technical report, UCL, 2016.

Xavier Giroud and Holger M. Mueller. Firms□ Internal Networks and Local Economic Shocks. Technical report, MIT, 2017.

Bryan S. Graham. An Empirical Model of Link Formation with Degree Heterogeneity. *Econometrica*, forthcoming.

Michael Greenstone, Richard Hornbeck, and Enrico Moretti. Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, 118(3):536–598, 2010.

Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic Blockmodels: First Steps. *Social networks*, 5(2):109–137, 1983.

Lawrence Hubert and Phipps Arabie. Comparing Partitions. *Journal of Classification*, 2(1): 193–218, 1985.

Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

Matthew O. Jackson and Brian W. Rogers. The Economics of Small Worlds. *Journal of the European Economic Association*, 3(2-3):617–627, 2005.

Brian Karrer and Mark E.J. Newman. Stochastic Blockmodels and Community Structure in Networks. *Physical Review E*, 83(1):016107, 2011.

Rafael Lalive. How do Extended Benefits Affect Unemployment Duration? A Regression Discontinuity Approach. *Journal of Econometrics*, 142(2):785 – 806, 2008.

Rafael Lalive and Josef Zweimüller. Benefit Entitlement and Unemployment Duration: The Role of Policy Endogeneity. *Journal of Public Economics*, 88(12):2587 – 2616, 2004.

Rafael Lalive, Camille Landais, and Josef Zweimüller. Market Externalities of Large Unemployment Extension Programs. *American Economic Review*, 105(12):3564–96, 2015.

Michael Lechner, Conny Wunsch, and Patrycja Scioch. Do Firms Benefit from Active Labour Market Policies? Economics Working Paper Series 1318, University of St. Gallen, 2013.

Alan Manning and Barbara Petrongolo. How Local are Labor Markets? Evidence from a Spatial Job Search Model. *American Economic Review*, forthcoming.

Enrico Moretti. Workers' Education, Spillovers, and Productivity: Evidence from Plant-level Production Functions. *American Economic Review*, 94(3):656–690, 2004.

Enrico Moretti. Local Labor Markets. In *Handbook of Labor Economics*, volume 4, pages 1237–1313. Elsevier, 2011.

Mark E.J. Newman and Michelle Girvan. Finding and Evaluating Community Structure in Networks. *Physical review E*, 69(2):026113, 2004.

Matthew J. Notowidigdo. The Incidence of Local Labor Demand Shocks. Technical report, National Bureau of Economic Research, 2013.

Tiago P Peixoto. Efficient Monte Carlo and Greedy Heuristic for the Inference of Stochastic Block Models. *Physical Review E*, 89(1):012804, 2014a.

Tiago P. Peixoto. Hierarchical Block Structures and High-resolution Model Selection in Large Networks. *Physical Review X*, 4(1):011047, 2014b.

Jennifer P. Poole. Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility. *Review of Economics and Statistics*, 95(2):393–406, 2013.

Ian M. Schmutte. Free to Move? A Network Analytic Approach for Learning the Limits to Job Mobility. *Labour Economics*, 29:49–61, 2014.

Michel Serafinelli. Good Firms, Worker Flows and Local Productivity. Technical report, University of Toronto, 2014.

Isaak Sorkin. Ranking Firms Using Revealed Preference. Technical report, Unpublished Manuscript, 2015.

Andrey Stoyanov and Nikolay Zubanov. Productivity Spillovers across Firms through Worker Mobility. *American Economic Journal: Applied Economics*, 4(2):168–198, 2012.

Andrey Stoyanov and Nikolay Zubanov. The Distribution of the Gains from Spillovers through Worker Mobility between Workers and Firms. *European Economic Review*, 70: 17–35, 2014.

Yuchung J. Wang and George Y. Wong. Stochastic Blockmodels for Directed Graphs. *Journal of the American Statistical Association*, 82(397):8–19, 1987.

Xiaoran Yan, Cosma Shalizi, Jacob E. Jensen, Florent Krzakala, Cristopher Moore, Lenka Zdeborova, Pan Zhang, and Yaojia Zhu. Model Selection for Degree-corrected Block Models. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(5):P05007, 2014.

Yunpeng Zhao, Elizaveta Levina, and Ji Zhu. Consistency of Community Detection in Networks under Degree-corrected Stochastic Block Models. *The Annals of Statistics*, 40(4): 2266–2292, 2012.

Josef Zweimüller, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philippe Wuellrich, Oliver Ruf, and Simon Büchi. Austrian Social Security Database. *NRN: The Austrian Center for Labor Economics and the Analysis of the Welfare State*, 2009.

1.A Network Definition and Characteristics

1.A.1 A formal definition of the job mobility network

Let $\mathcal{M}$ be the set of firms in the economy and $\mathcal{W}$ be the set of workers in the economy. Further, let $t \in \{1, \ldots, T\}$ denote the time in days during the sample period. I define a function $m = m(w, t) \in \mathcal{M}$ that returns the firm that employs worker w at time t. Then we have the following definition of the entries in the adjacency matrix of a directed and weighted network:

$$
\begin{aligned}
A_{ij} = |\omega| \quad \text{where} \quad \omega = \{w \mid\ & \exists\, t \in T,\ 1 \leq x \leq 30 \text{ s.th.} \\
& m(w, t-365) = \cdots = m(w, t) = i \\
& m(w, t+x) = m(w, t+x+1) = \cdots = m(w, t+x+365) = j \\
& m(w, t+1), \ldots, m(w, t+x-1) \notin \mathcal{M}\}.
\end{aligned}
$$

While the set of firms $\mathcal{M}$ contains all firms that exist in the economy during the sample period, the set of nodes in the network, N, contains only non isolates:

$$N = \mathcal{M} \setminus \{i \in \mathcal{M} \mid A_{ij} = 0\ \forall j\ \wedge\ A_{ji} = 0\ \forall j\}. \tag{1.13}$$

1.A.2 Empirical Network Characteristics

In this section, I provide additional information on the empirical characteristics of the job mobility network described in section 1.3. Table 1.A5 provides various characteristics of the job mobility network. Again, column 1 refers to the full network obtained from transitions between 1975 and 2005 while the other columns show dynamic developments between shorter periods.

Panel A of Table 1.A5 displays the most important network characteristics that describe the general structure of the job mobility network. The *average degree* in the network denotes the average number of transitions per firm. On average, a firm is connected to 19.84 other firms in the full network. The average degree is naturally lower when considering shorter time periods and tends to slightly rise over time. The ease of information flows in a network can be measured by the notion of distance between nodes. In the job mobility network, the shortest path between the two most distant firms (called *diameter* of the network) requires 17 steps (row 2). The average number of steps along the shortest paths between all possible pairs of firms in the network amounts to 4.87 steps (*average path length* in row 3). The density of the job mobility network is very sparse as only a tiny fraction of all possible links materializes (*graph density* in row 4). Finally, the *clustering coefficient* (row 5) measures the transitivity of a network, i.e., the probability that two firms with common links to a third firm are linked among themselves.

Transitivity in the job mobility network is relatively low compared to other social networks as only 2-4% of all potential triangles materialize.[41] This finding suggests that, on the very local level of three firms, job-to-job mobility is not particularly clustered. In the descriptive analysis in section 1.5, I specifically show that job-to-job transitions are clustered on a broader labor market level rather than on subsets of a few firms.

Aggregated network characteristics such as the average degree potentially hide substantial heterogeneity within the network. In the present case, many firms in the job mobility network are involved only in a low number of job-to-job transitions while others have many connections and serve as "hubs" in the economy. This is documented by the (complementary) CDF of the degree distribution in Figure 1.A14. The black circles represent the empirical CDF of the degree distribution on a log-log scale. Like in many social networks, the degree distribution of the job mobility network exhibits heavy tails, as there are more nodes with very small and very large degrees than expected in a model where links are formed uniformly at random (Jackson, 2008).

In an influential paper, Jackson and Rogers (2005) analyze the interdependence between the process of link formation in social networks and the degree distribution. In a nutshell, a model where new nodes form links to existing ones *uniformly at random* is consistent with an exponential degree distribution. In contrast, a model of *preferential attachment*, where the probability to receive links for existing firms is proportional to their current degree, is consistent with a degree distribution that follows a power law.

The colored lines in Figure 1.A14 therefore show maximum likelihood fits from both, the exponential distribution (in blue) and the power law distribution (in red). The parameter estimates of the fitted distributions are given in the first two rows of panel B in Table 1.A5. Neither of these distributions is a good fit for the degree distribution of the job mobility network.

Most empirical networks are somewhere in between the extreme cases of random link formation and preferential attachment. Jackson and Rogers (2005) therefore develop a hybrid model where a fraction r of links is formed uniformly at random while the remainder is generated based on preferential attachment. The green line in Figure 1.A14 displays the fit of this hybrid model which is much closer to the observed degree distribution. The estimate in the third row of panel B in Table 1.2 indicates that 39% of links in the job mobility network are formed uniformly at random while the majority of 61% are formed through network-based link generation.[42]

[41] The clustering coefficient in the job mobility network is higher than it would be if links were formed purely random (19.84/95237 = 0.0002). However, Jackson (2008) reports much higher coefficients obtained from various other social networks.

[42] Figure 1.A15 illustrates the degree distribution and parametric fits for the shorter time periods. Although the share of random and network based link formation varies to some extend, the general picture is very stable over time.

Summing up, there is strong evidence for preferential attachment in the link formation process. In particular, workers tend to join firms that received an influx of many other workers and leave firms that are left by many others. Although not taking a dynamic perspective on link formation, I specifically address this form of firm-level heterogeneity in the model for the estimation of endogenous labor markets in section 1.4 by including popularity parameters that guide the individual attractiveness of firms to workers.

1.B Additional Derivations

If match-quality draws are distributed according to the Frechét distribution, the probability that worker ℓ starting in firm i has a payoff higher than some threshold φ in firm j is

$$Pr[\varphi_j(\ell|i) > \varphi] = 1 - F_j\left(\frac{\varphi d_{z_i z_j}}{p_j f[X_\ell]}\right) = 1 - e^{-T_j d_{z_i z_j}^{-\theta}(p_j f[X_\ell])^{\theta}\phi^{-\theta}} \tag{1.14}$$

and the probability that the payoff is lower than φ in all other firms $s \neq j$ is

$$Pr[\varphi_s(\ell|i) \leq \varphi, \forall s \neq j] = \prod_{s\neq j} F_s\left(\frac{\varphi d_{z_i z_s}}{p_s f[X_\ell]}\right) = \prod_{s\neq j} e^{-T_s d_{z_i z_s}^{-\theta}(p_s f[X_\ell])^{\theta}\phi^{-\theta}} \tag{1.15}$$

As a result, the probability that j offers the highest payoff of all firms for a worker ℓ who starts in i is

$$\begin{aligned}
\pi_{ij}(\ell) &= Pr[\varphi_j(\ell|i) \geq \max_s\{\varphi_s(\ell|i)\}] \\
&= \int_0^\infty Pr[\varphi_s(\ell|i) \leq \varphi, \forall s \neq j] dPr[\varphi_j(\ell|i) \leq \varphi] \\
&= \frac{T_j d_{z_i z_j}^{-\theta} p_j^{\theta}}{\sum_{s=1}^N T_s d_{z_i z_s}^{-\theta} p_s^{\theta}}
\end{aligned}$$

1.C Simulation

In order to evaluate the performance of estimating the degree-corrected stochastic block model, I conduct a Monte-Carlo simulation exercise that compares the SBM to using predefined regions in a stylised economy.

The economy consists of N firms $i = \{1, \ldots, N\}$ that are located in either of two regions $r_i \in \{1, 2\}$. With probability τ a firm resides in region 1 and with probability $1 - \tau$ it resides in region 2. Hence, varying the parameter τ allows to examine the robustness of the model to changes in the relative size of the regions. The actual market assignments, however, are governed by an "unobserved" characteristic z which can be correlated with the region membership. The unobserved characteristic $z_i \in \{1, 2\}$ can take two distinct values and is distributed conditional on the region membership as follows: $P(z_i = r_i) = \lambda$, $P(z_i \neq r_i) = 1-\lambda$. Hence, varying the parameter λ from 1 (perfect positive correlation) to 0 (perfect negative correlation) determines in how far region membership guides actual market assignment. The firms in the economy are furthermore characterised by degree parameters γ_i. Degrees are drawn from a power law distribution with minimum expected degree of x_{min} and parameter α. The degree parameters γ_i are then determined fixed according to equation (1.8). The transition matrix between the markets is set to

$$M = \varrho \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix},$$

where ϱ is chosen such that it fixes the overall expected degree of the network. Finally, links between firms i and j in the economy are drawn from the Poisson distribution with mean $\gamma_i \gamma_j M_{z_i z_j}$.

The parameters in the simulation study are chosen as follows: There are $N = 1000$ firms. The group sizes are balanced in a version with $\tau = 0.5$ and unbalanced in a version with $\tau = 0.75$. The power law resembles the actual degree distribution found in the Austrian job mobility network with $x_{min} = 20$ and $\alpha = 2.5$. Similarly, ϱ is chosen such that the overall average degree equals 8 as in the empirical network. To compare the solution of estimating the SBM to the true assignments and to the use of the region membership, I use the adjusted Rand index of Hubert and Arabie (1985) and the normalized mutual information criterion of Danon, Diaz-Guilera, Duch, and Arenas (2005). These indices are commonly used to measure the similarity between partitions in clustering and network analysis. Both measures are scaled such that 1 corresponds to a perfect match between two partitions while a value of 0 zero would be expected for two random partitions.

Figure 1.A16 displays the median adjusted Rand index over 100 replications varying the correlation coefficient between regions and group assignments from 1 to 0. In panel a, the group sizes are balanced. As expected, the concordance of the predefined regions with the true group assignments decreases with a declining correlation between the two random variables. When λ equals 0.5 group assignments are independent from region membership and the Rand index approaches 0. In contrast, the degree-corrected SBM does not depend on the region membership and therefore constantly achieves high scores of the adjusted Rand index which are of similar magnitude as the simulation results for sparse networks in Zhao et al. (2012). The results in an unbalanced setting (panel b) are very similar. The estimation of the SBM, however, is a bit less precise as indicated by the standard error bars.

The results of this simulation study indicate that even for slight deviations from perfect congruence of regions and relevant labor markets, it is favorable to base the analysis on the degree-corrected SBM proposed in this paper. The fact that the SBM does not rely on observed covariates but infers the group structure solely based on observed links enables a stable detection of relationships independent of whether the relevant covariates are known or available.

1.D Additional Tables and Figures

Table 1.A1: Transition Probabilities and Market Characteristics

Market ID	1	2	3	4	5	6	7	8	9
Panel A. Transition Probabilities									
1	.096	.001	.017	.001	.003	.002	.001	.006	.010
2	.001	.059	.001	.001	.005	.001	.004	0	0
3	.019	.001	.12	.001	.002	.003	.001	.015	.012
4	.001	.001	.001	.111	.001	.001	.002	.001	.001
5	.003	.006	.002	.001	.13	.001	.009	.001	.001
6	.003	.001	.002	.001	.002	.083	.006	.001	.001
7	.001	.002	.001	.002	.003	.003	.074	.001	.003
8	.005	.001	.01	.001	.001	.001	.001	.06	.002
9	.003	0	.004	.001	.001	.001	.002	.002	.057
Panel B. Market Characteristics									
Mode state	NOE	Stm.	Wien	Tirol	OOE	Szbg	Ktn	Wien	Wien
Share in mode state	0.56	0.80	0.63	0.58	0.91	0.80	0.76	0.69	0.48
Mode Industries	Constr. Retail Business	Retail Constr. Hot.& Rest.	Wholesale Retail Constr.	Constr. Retail Wholesale	Constr. Retail Wholesale	Constr. Retail Wholesale	Retail Constr. Health	Business Wholesale Finance	Health Publ. Adm. Constr.
Share in mode industry	0.18	0.14	0.13	0.09	0.10	0.09	0.09	0.22	0.23
Avg. firm size	31.738 (108.02)	28.910 (108.98)	30.966 (100.32)	32.736 (126.03)	38.341 (262.15)	35.468 (162.31)	46.051 (327.12)	58.203 (265.10)	101.168 (910.68)
Observations	15666	7890	17439	12332	14535	12401	6618	3431	3324

Note: Standard deviations in parenthesis. Panel a. of this table presents the estimated transition probability matrix $\hat{M}$ from the SBM indicating transition probabilites within and between markets. Panel b. reports market characteristics for each market. Note that for some firms industry information is missing. The entries in the 3rd row refer to the three most common non-missing industries.

Table 1.A2: Industry Composition of Endogenous Steel Market outside Linz-Wels

NACE 3-digit industry		Share of employment
452	Building of complete constructions	7.30
341	Manufacture of motor vehicles	6.95
361	Manufacture of furniture	5.66
751	Administration of the State	3.66
247	Manufacture of man-made fibres	3.24
524	Retail sale in specialized stores	3.24
287	Manufacture of other fabricated metal products	3.21
182	Manufacture of wearing apparel and accessories	2.59
602	Other land transportation	2.22
453	Building installation	1.96
293	Manufacture of agricultural and forestry machinery	1.91
193	Manufacture of footware	1.84
502	Maintainance and repair of motor vehicles	1.76
651	Monetary intermediation	1.73
211	Manufacture of pulp, paper amd paperboard	1.67
159	Manufacture of beverages	1.55
295	Manufacture of other special purpose machinery	1.54

Note: This table reports NACE 3-digit industry affiliations of the firms in endogenous labor market of the steel company but outside the NUTS3 region Linz-Wels for the years 1980-1990. The share for each industry is weighted by employment.

Table 1.A3: Industries with Highest Increase in Imports from the East over 1990 to 2010

NACE 3-digit industry		Percent increase
283	Steam generators	183.96
354	Motorcycles and bicycles	160.23
233	Nuclear fuel	63.10
341	Motor vehicles	51.43
243	Paints, coatings, printing ink	48.76
322	TV, and radio transmitters, apparatus for line telephony	46.46
312	Electricity distribution and control apparatus	46.03
267	Cutting, shaping, finishing of stone	40.08
176	Knitted and crocheted fabrics	38.55
222	Printing	38.02
245	Detergents, cleaning and polishing, perfumes	30.98
273	Other first processing of iron and steel	19.79
343	Parts and accessories for motor vehicles	19.66
221	Publishing	18.45
282	Tanks, reservoirs, central heating radiators and boilers	17.82
291	Machinery for production, use of mech. power	16.99
313	Isolated wire and cable	15.46
183	Dressing and dyeing of fur; articles of fur	14.89
314	Accumulators, primary cells and primary batteries	14.78
295	Other special purpose machinery	12.10
268	Other non-metallic mineral products	12.03
342	Bodies for motor vehicles, trailers	11.85
316	Electrical equipment n. e. c.	11.82
334	Optical instruments and photographic equipment	11.75
177	Knitted and crocheted articles	11.50

Table 1.A4: Industries with Highest Increase in Exports to the East over 1990 to 2010

NACE 3-digit industry		Percent increase
202	Panels and boards of wood	62.57
296	Weapons and ammunition	29.56
233	Nuclear fuel	24.52
204	Wooden containers	21.83
265	Cement, lime and plaster	18.93
153	Fruits and vegetables	15.71
264	Bricks, tiles and construction products	15.01
171	Textile fibres	12.60
353	Aircraft and spacecraft	10.36
191	Tanning and dressing of leather	10.28
151	Meat products	9.70
172	Textile weaving	8.32
176	Knitted and crocheted fabrics	7.92
342	Bodies for motor vehicles, trailers	7.52
341	Motor vehicles	6.46
181	Leather clothes	6.14
354	Motorcycles and bicycles	6.02
293	Agricultural and forestry machinery	5.89
334	Optical instruments and photographic equipment	5.78
343	Parts and accessories for motor vehicles	5.40
314	Accumulators, primary cells and primary batteries	4.26
183	Dressing and dyeing of fur; articles of fur	4.11
274	Basic precious and non-ferrous metals	3.87
192	Luggage, handbags, saddlery and harness	3.64
193	Footwear	3.59

Table 1.A5: Network Characteristics of the Job Mobility Networks

	1975-2005	1975-1980	1980-1985	1985-1990	1990-1995	1995-2000	2000-2005
Panel A. Network Characteristics							
average degree	19.84	9.98	8.42	9.00	9.59	9.31	9.75
diameter	17	20	20	19	20	25	19
average path length	4.87	5.67	5.83	5.74	5.74	5.95	5.95
density ($\times 10,000$)	1.06	0.97	0.88	0.87	0.85	0.84	0.86
clustering coefficient	0.04	0.03	0.02	0.03	0.03	0.03	0.03
Panel B. Degree Distribution							
Power Law Dist. α	2.56	2.56	2.53	2.43	2.48	2.57	2.43
Exponential Dist. λ	0.34	0.43	0.47	0.45	0.44	0.46	0.45
Hybrid Model fraction r	0.37	0.74	0.57	0.59	0.62	0.7	0.5

Note: All measures correspond to the giant component of the job mobility network sampled during the years indicated. Avg. degree measures the average number of incoming and outgoing connections per firm. The diameter is the shortest path between the two most distant firms in the network. Avg. path length indicates how many steps on average it takes to get from one firm to another. Graph density is the fraction of all possible links that are actually present. The clustering coefficient measures the fraction of actually observed triangles in all potential triangles in the network. Panel B reports parameters from parametric fits to the degree distribution for three different models.

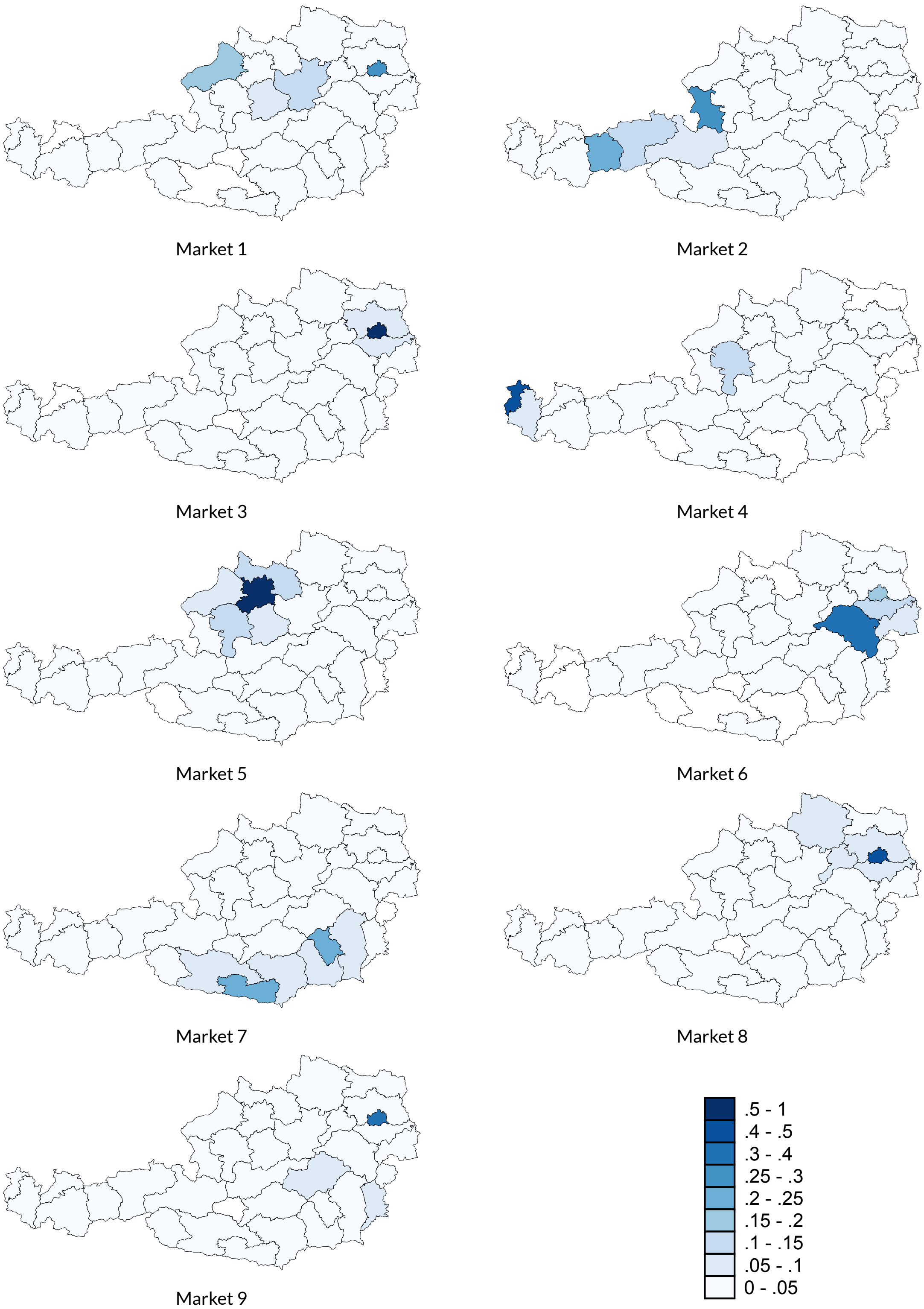


Figure 1.A1: Share of Firms in NUTS-3 Regions for each Market (1975-1980)

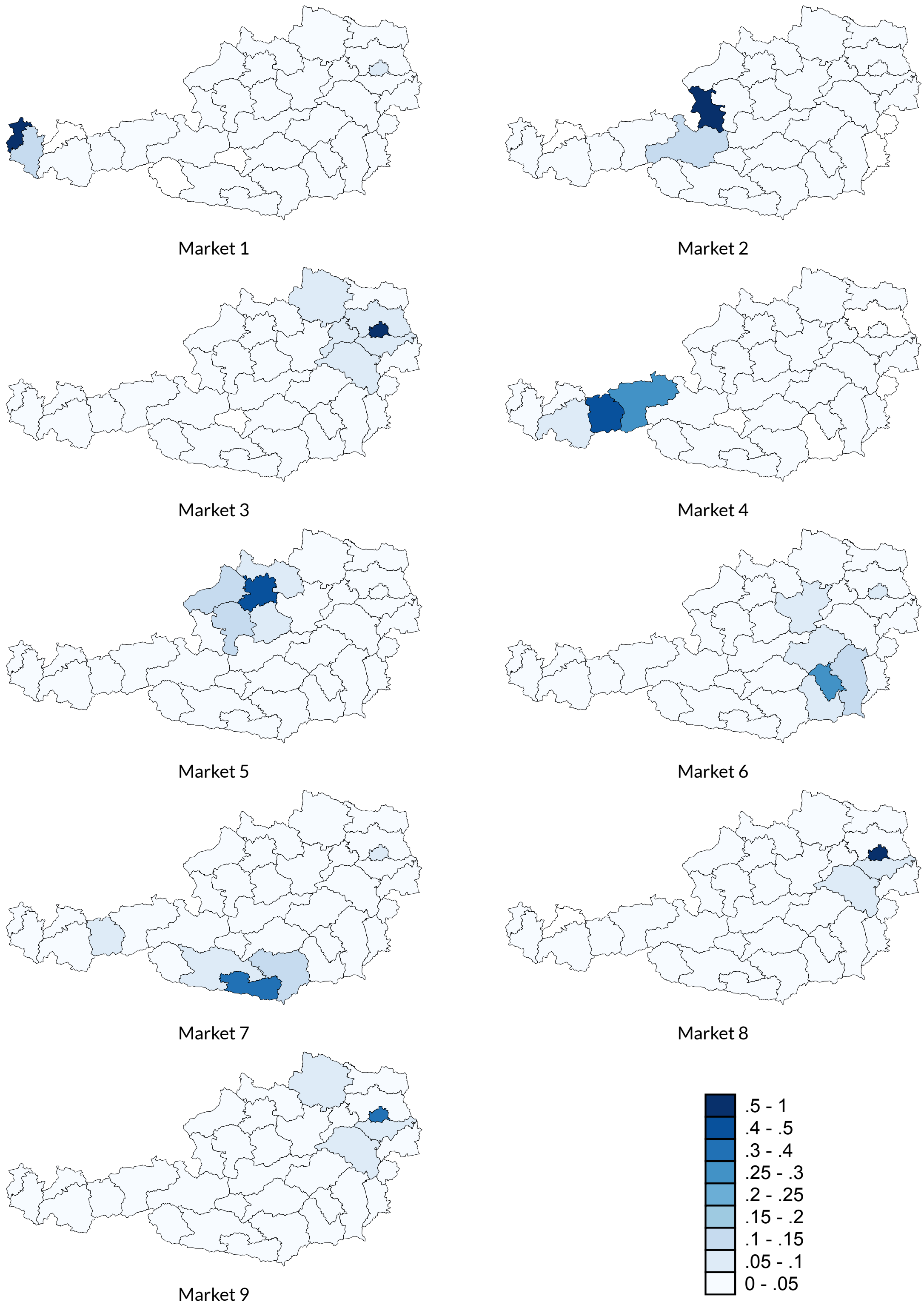


Figure 1.A2: Share of Firms in NUTS-3 Regions for each Market (1980-1985)

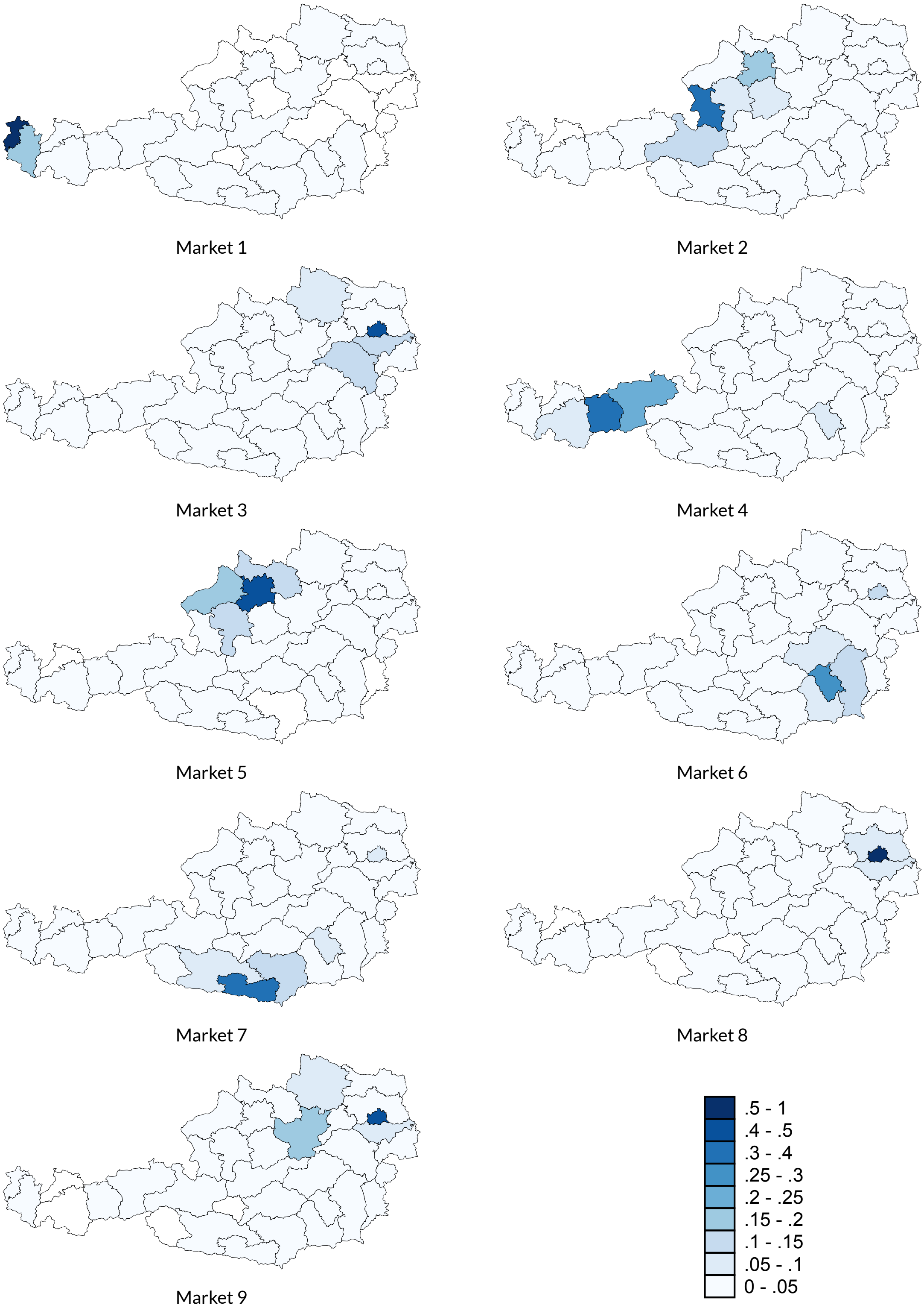


Figure 1.A3: Share of Firms in NUTS-3 Regions for each Market (1985-1990)

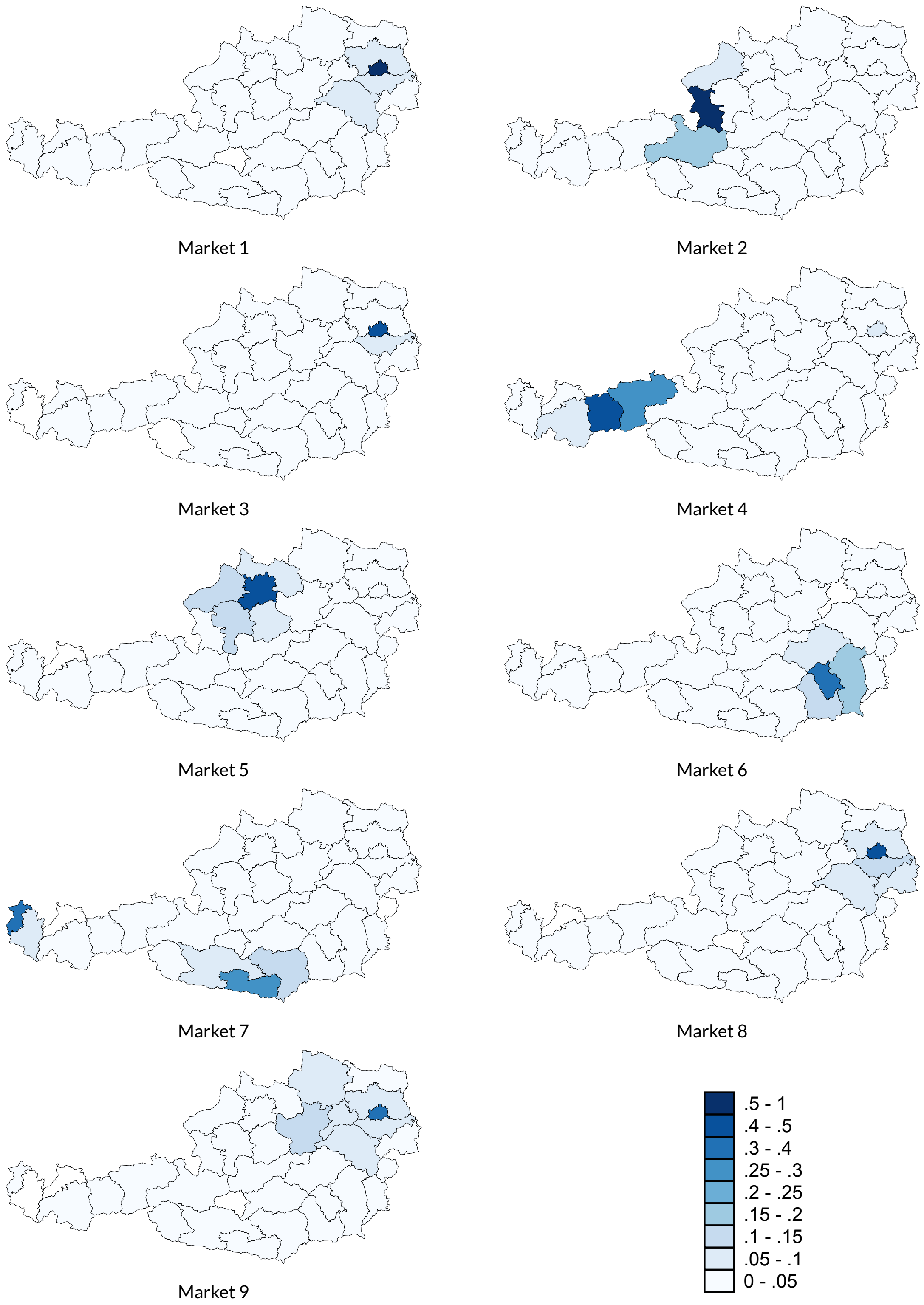


Figure 1.A4: Share of Firms in NUTS-3 Regions for each Market (1990-1995)

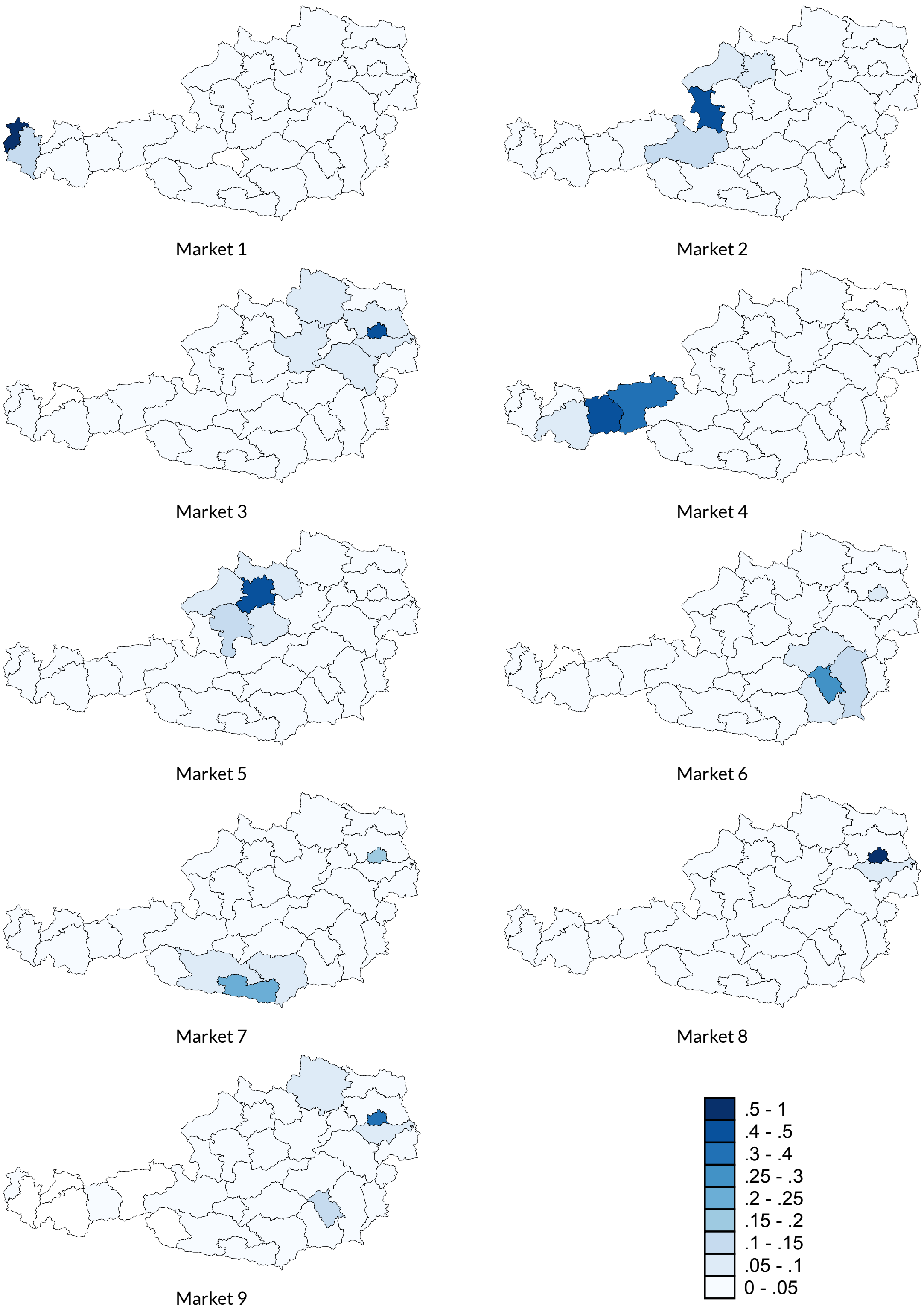


Figure 1.A5: Share of Firms in NUTS-3 Regions for each Market (1995-2000)

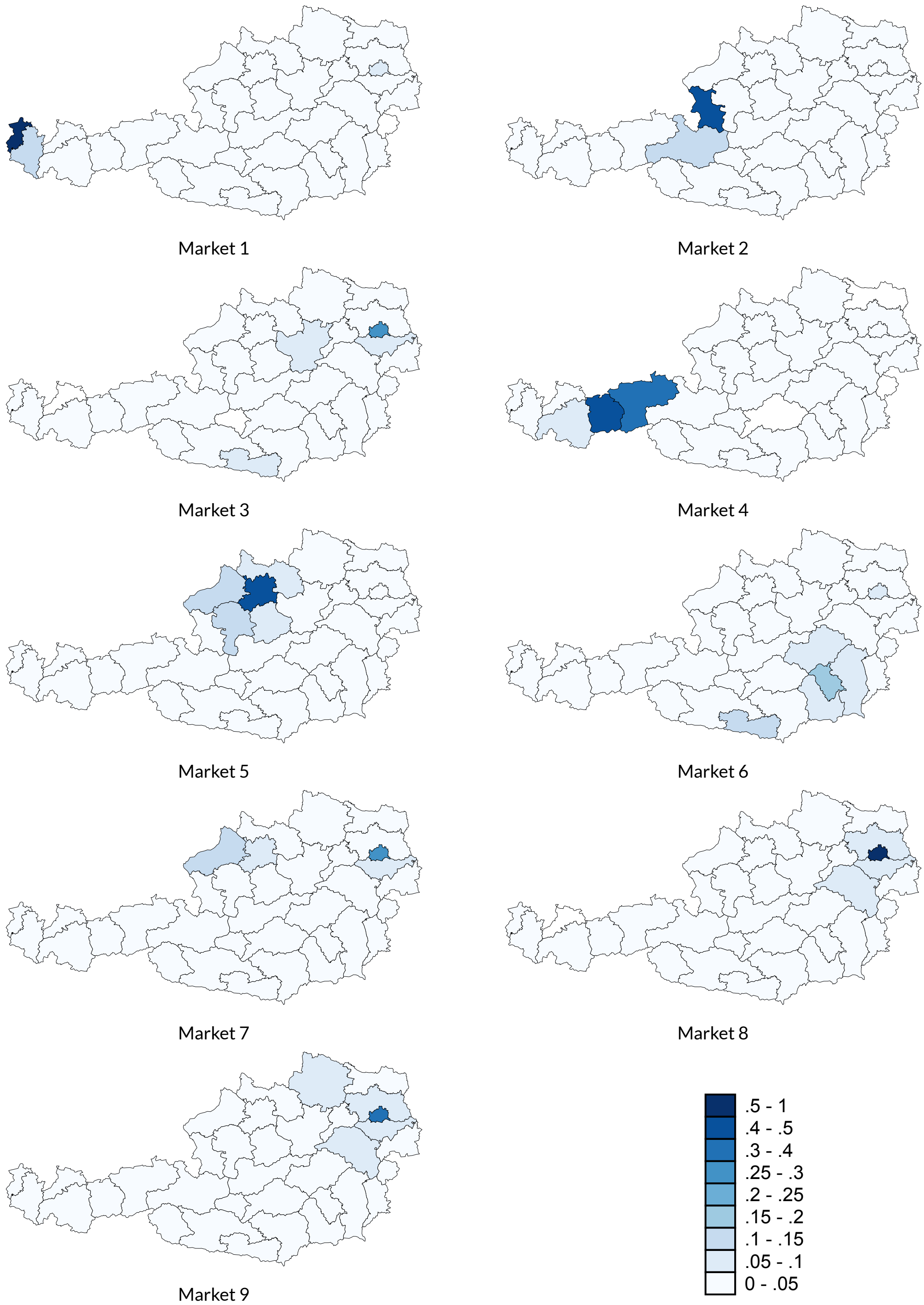


Figure 1.A6: Share of Firms in NUTS-3 Regions for each Market (2000-2005)

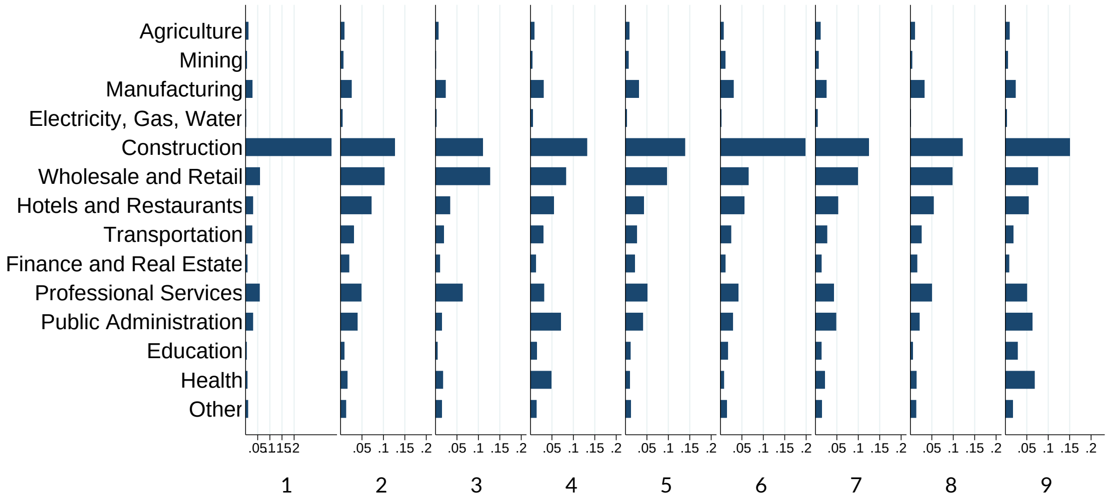


Figure 1.A7: Histogram of Industry Composition by Market (1975-1980)

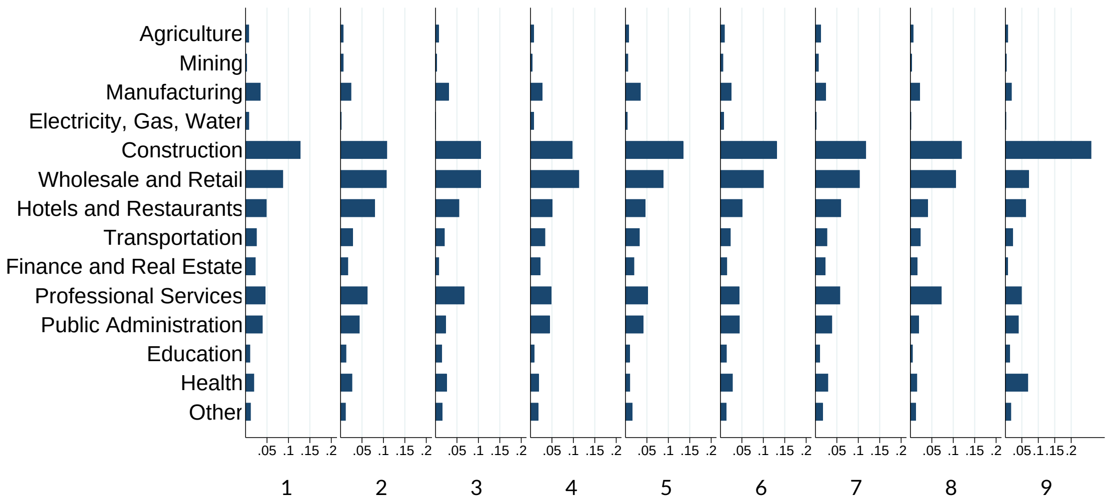


Figure 1.A8: Histogram of Industry Composition by Market (1980-1985)

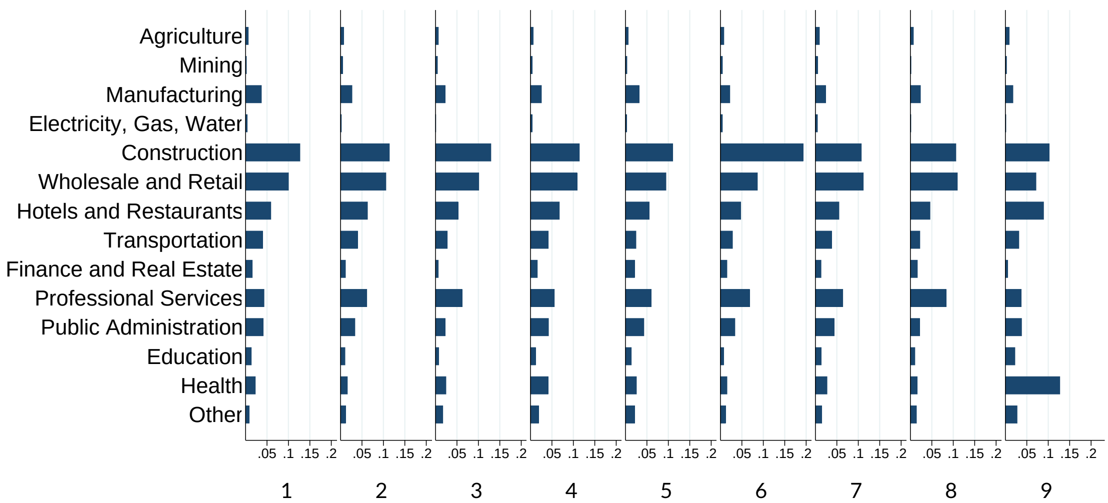


Figure 1.A9: Histogram of Industry Composition by Market (1985-1990)

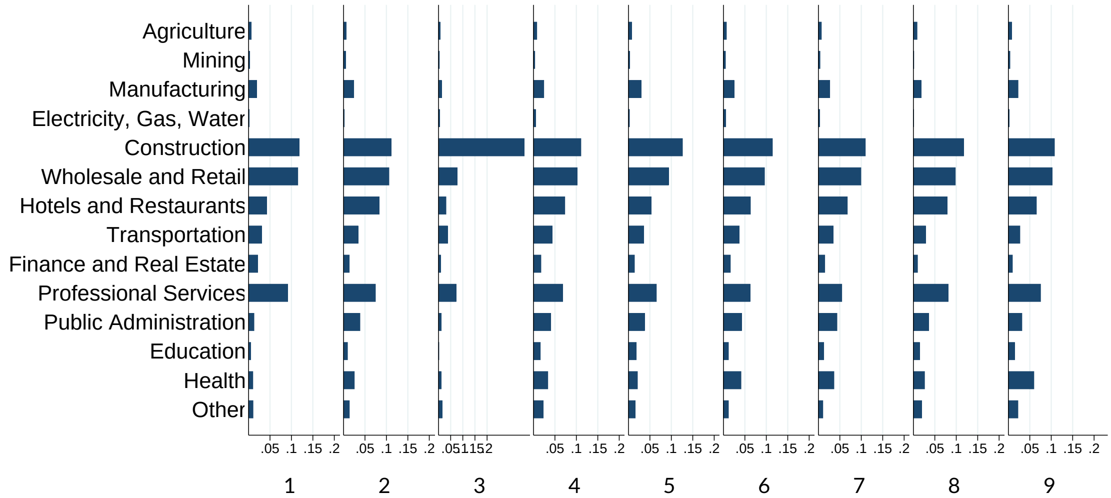


Figure 1.A10: Histogram of Industry Composition by Market (1990-1995)

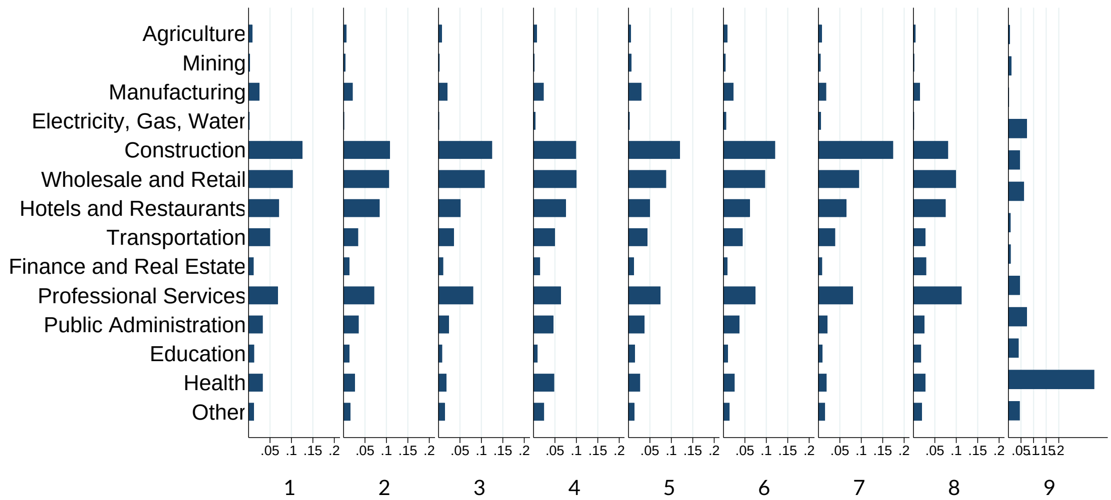


Figure 1.A11: Histogram of Industry Composition by Market (1995-2000)

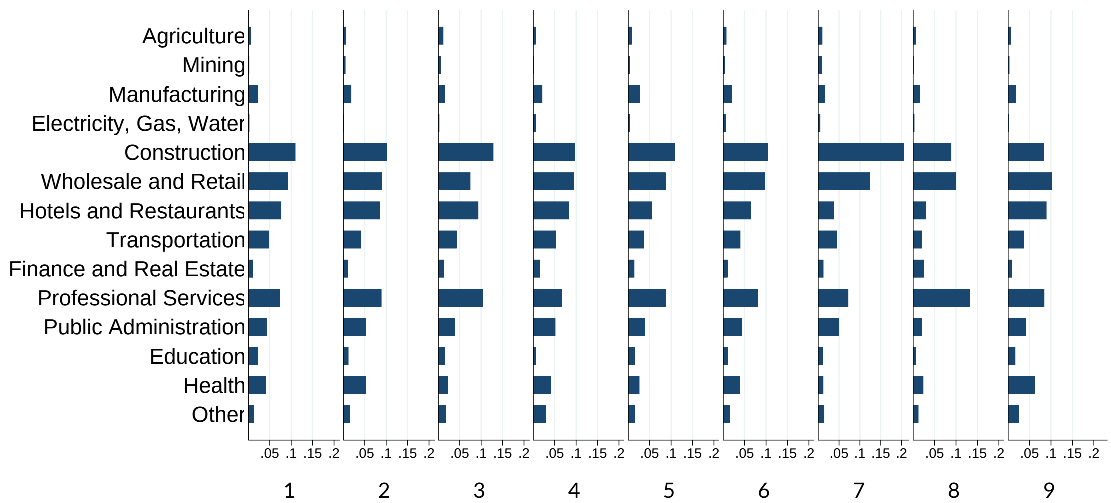


Figure 1.A12: Histogram of Industry Composition by Market (2000-2005)

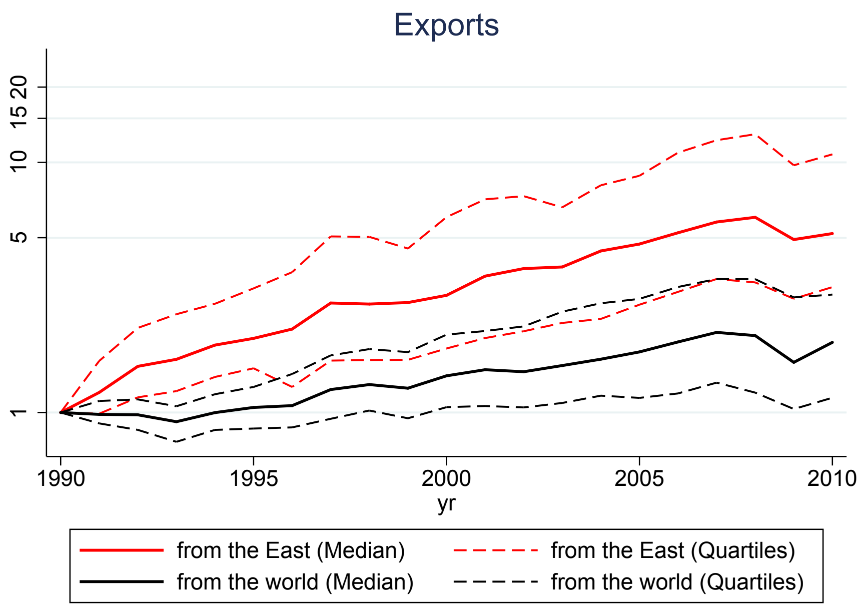


Figure 1.A13: Rising Export Volumes in Austrian Trade

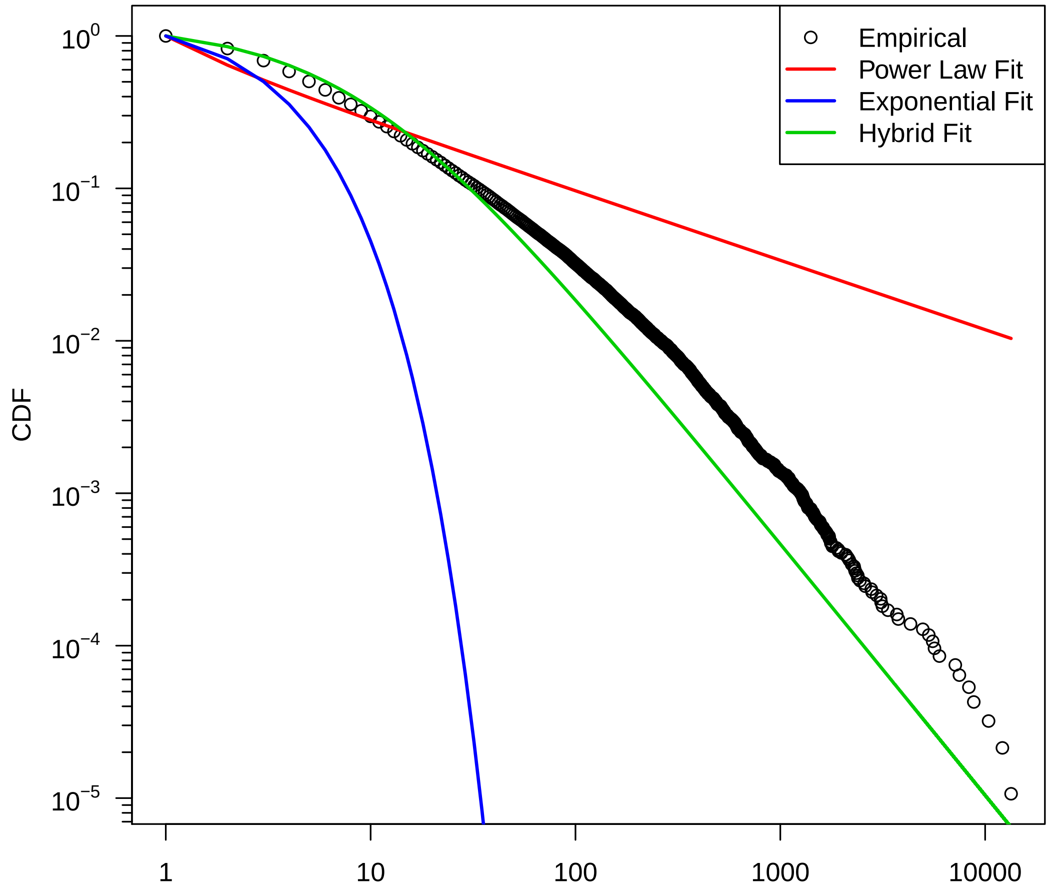


Figure 1.A14: Complementary CDF of the Degree Distribution in the Job Mobility Network 1975-2005

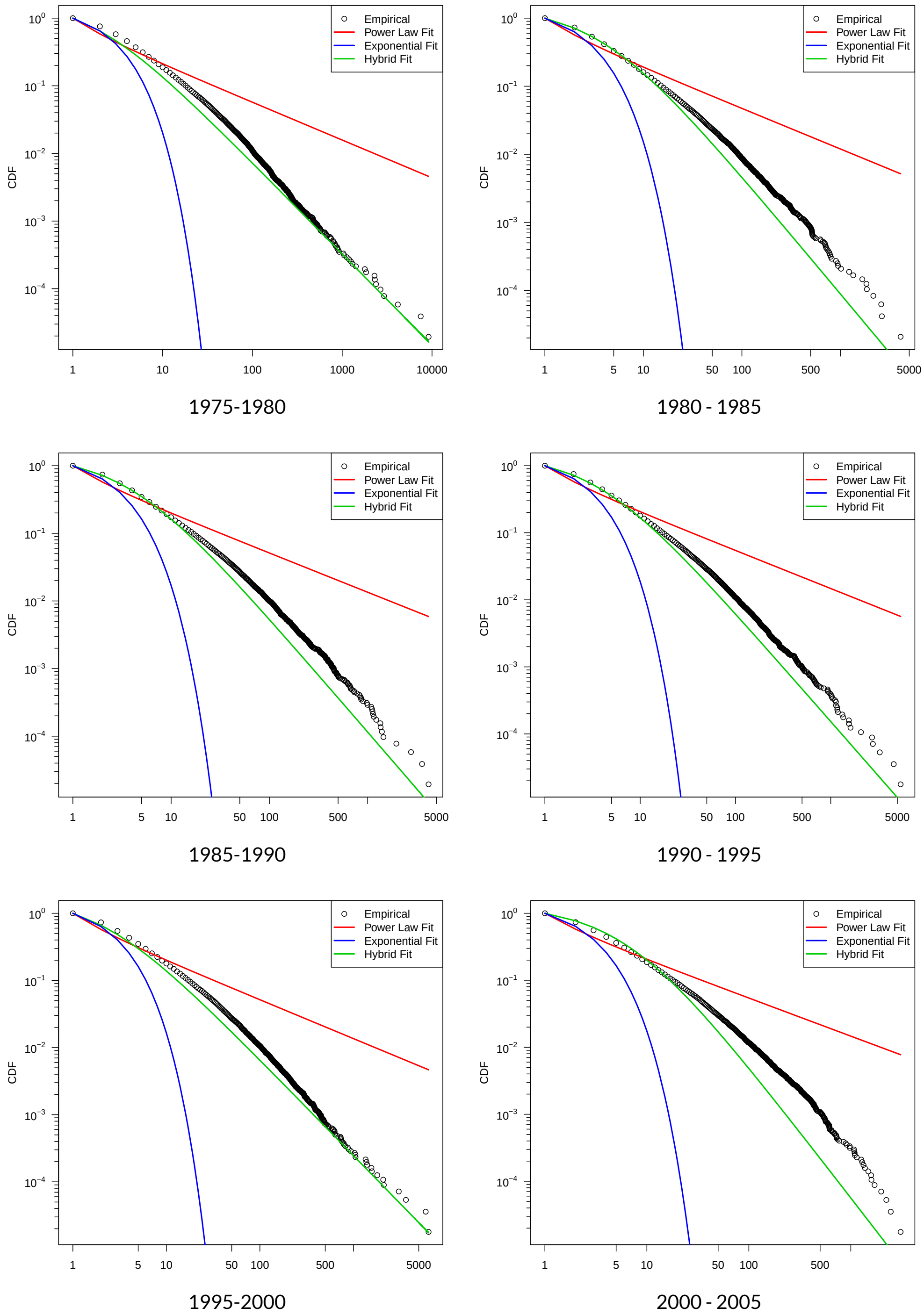
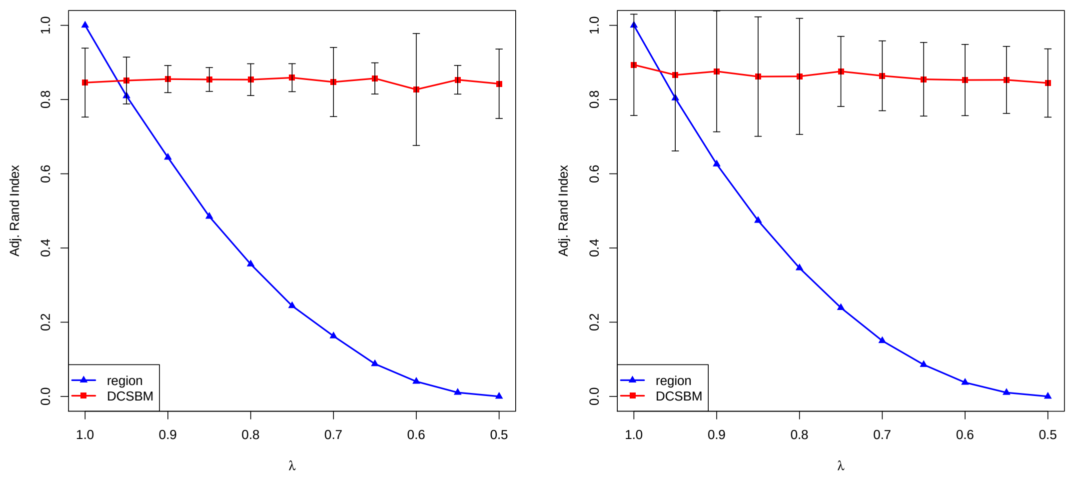


Figure 1.A15: Complementary CDF of the Degree Distribution in the Job Mobility Network over Time



(a) Balanced Group Size (τ = 0.5)

(b) Unbalanced Group Size (τ = 0.75)

Figure 1.A16: Results for the Predefined Regions and the Degree-corrected SBM, λ varies, error bars indicate one standard deviation to each side

Chapter 2

Learning Dynamics in Tax Bunching at the Kink: Evidence from Ecuador

with Albrecht Bohne

2.1 Introduction

Despite the predictions of labor supply models, empirical studies have only found limited evidence for bunching behavior at kink points in the marginal tax schedule. Information frictions are a commonly used explanation for the absence of pronounced spikes in the income distribution in the literature on behavioral responses to taxation. An important open question is how individuals that are new to the institutional setting of paying taxes react to incentives posed by the system. Moreover, there is no clear consensus on how information about tax adjustment opportunities is transmitted and what the driving factors of these adjustments are.

In this paper, we exploit new and very detailed administrative data on personal income tax (PIT) returns in a developing country, Ecuador. The environment of a rapidly formalizing economy with a steady inflow of new individuals to the tax system provides a unique setting to study the dynamics of tax responses. This is especially relevant since in the process of formalization, developing countries rely ever more on PIT (Besley and Persson, 2013). Particularly, we examine how workers' responses to jumps in the marginal tax rate (inducing kinks in their budget sets) change over time and with increasing experience and exposure to the tax system. Furthermore, we contribute to the literature by disentangling the effects that firm-level practices and co-worker behavior have on individual tax-filing.

Using new individual tax return data on the universe of formal-sector wage earners in Ecuador ranging from 2006 to 2015, we provide evidence for substantial sensitivity of reported taxable income to a discontinuous jump in the marginal tax rate. We observe a large and pronounced spike in the distribution of taxable income just before the income tax exemption threshold. We quantify the prevalence of this bunching behavior using an established bunching estimator which relates the excess mass in this area to an estimated counterfactual (Kleven, 2016). The effect is primarily driven by about 20% of the working population who take advantage of generous deduction possibilities in health, education, housing, clothing, and food. These deduction possibilities are the main part of the Ecuadorian government's policies to induce an increase in formalization. The tax responses shown in the data represent reporting behavior rather than real labor supply responses as there is no indication of bunching in gross income. Most importantly, the mass of bunchers in taxable income increases with higher experience in filing taxes and stronger exposure to the tax incentives. This leads to the conclusion that workers in Ecuador learn about the incentives and measures to avoid paying taxes as they adjust to the system.

However, it is unclear how exactly workers learn about the tax system. We shed light on how individuals learn about these tax adjustment opportunities and what the predominant channels of information transmission are. Based on detailed matched employer-employee data and a research design that exploits job transitions, we can disentangle whether the observed learning patterns are mainly driven by individuals learning from firms (and firm-level institutions) or individuals learning from their co-workers.

To quantify how individuals learn about tax adjustment opportunities from their firm, we generate a sample of job switchers who change their main employer within our sample period and track the degree of bunching among their co-workers in the old and new firms. Our results show a strong and asymmetric adjustment to the prevailing bunching practices at the firm level. The probability to bunch for individuals who move to a firm in the top quintile of the distribution of bunching shares (from an origin firm in the middle quintile) increases overall by about 3-5 percentage points while it remains constant when moving to a firm in the bottom quintile, even when controlling for a range of individual and firm-level characteristics. We show that the effects are persistent and even increase their magnitude in the second year at the new firm. The asymmetry of the effects lends strong support to the hypothesis that knowledge spillovers and memory play an important role in determining individual tax-filing behavior. Particularly, our evidence is consistent with a model of learning and memory in which individuals learn about tax adjustment opportunities when moving into a high-knowledge environment. When moving to a low-knowlege environment, however, individuals retain their previous knowledge and maintain their behavior with respect to taxes.

To shed light on the second possible learning mechanism at work, namely individuals learning from their co-workers, we generate a sample of firms that hire new employees. We compare bunching among incumbent employees in firms with incoming workers who were previously bunching to incumbent employees in firms with incoming workers who were not bunching before.[1] We find no evidence of workers learning about tax adjustment opportunities from new co-workers. Even among small firms and firms without any experience in bunching where we would expect larger effects, we cannot provide evidence for spillovers of new co-workers to incumbent workers. We conclude that in the setting of this study, firms seem to be a much stronger driver of individual tax-filing behavior than a given employee's co-workers.

The remainder of the paper is organized as follows. In section 2.2, we give an overview of the related literature and the contributions of this paper. Section 2.3 provides information on the institutional background in Ecuador and describes the PIT system in detail. Section 2.4 gives detailed information on the various data sources employed in our study. In section 2.5 we present the results from our analysis. Section 2.6 concludes.

2.2 Related Literature

Our paper contributes to the growing literature on bunching at kinks and notches in the tax schedule that was started by the seminal paper by Saez (2010). The method of estimating labor supply responses from the size of the excess mass at kinks and notches was further developed by Chetty, Friedman, Olsen, and Pistaferri (2011) and Kleven and Waseem (2013) and is thoroughly summarized in Kleven (2016). Evidence on behavioral responses to personal income taxation stems mainly from developed countries. Chetty et al. (2011) and Bastani and Selin (2014) analyze data from Scandinavia. They find bunching only at selected, particularly salient kinks (e.g., the top tax bracket) and for subgroups that can adjust their income relatively easily such as self-employed workers. In comparison, our results indicate relatively strong reactions to a very small kink.[2] Moreover, we concentrate on bunching solely among wage earners. In line with large parts of the literature, we find bunching to be driven mainly through reporting behavior and not real labor supply responses. The generous deduction possibilities in Ecuador are an interesting environment to study in this regard since they lend workers considerable scope to adjust their reported income.

Evidence on knowledge diffusion and spillover effects in bunching is provided by Chetty and Saez (2013), Chetty, Friedman, and Saez (2013), and Paetzold and Winner (2016). These papers analyze the effect of moving to high- or low-bunching environments and find significant impacts of coworker/regional bunching shares on individual bunching. Moreover, their evidence is supportive of learning and memory as individuals increase bunching when exposed

[1] We only regard incoming workers with previous gross income in the range where bunching at the first kink would have been possible.

[2] The first kink in the Ecuadorian tax schedule is very salient. The change in marginal tax rates from zero to five percent, however, is very small in international comparison.

to high-bunching environments but keep bunching when moving into low-bunching environments.

Our contribution to this literature is threefold. First, we add the dimension of experience with the tax system and find important impacts of previous exposure to the system on the adjustment process. Second, we analyze bunching of job switchers on the firm level and find much stronger effects than the studies that examine aggregate effects on the regional level. Third, we disentangle learning effects at the firm-level from those occuring between co-workers.

Another related strand of the literature is concerned with behavioral responses to taxes in developing countries. Kleven and Waseem (2013) and Best, Brockmeyer, Kleven, Spinnewijn, and Waseem (2015) analyze responses to notches in the PIT in Pakistan. Bachas and Soto (2015), Carrillo, Emran, and Rivadeneira (2012), and Carrillo, Pomeranz, and Singhal (2017) assess the reactions to incentives in corporate taxation. Most interestingly, the last two papers refer to data in Ecuador and find substantial evidence for tax evasion of firms in the country. As shown by Besley and Persson (2013), however, in the course of their progress towards more formal economies, developing countries rely increasingly on PIT. This lends importance and relevancy to our analysis of a personal income tax system in a developing country in the middle of this transition.

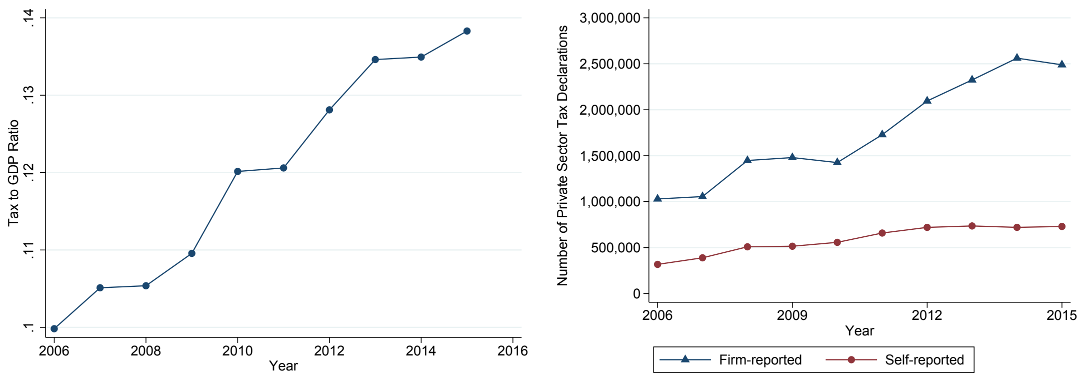
2.3 Institutional Background

Since 2008, Ecuador has implemented a wide range of economic and political reforms. The government has greatly increased spending on social programs and public service delivery. While a surge in oil revenues facilitated some of this increased spending, the tax administration has also pushed wide-ranging reforms of the tax system and tax collection policies. As a result, tax revenue as well as the tax base have grown substantially over the past years. Moreover, there has been a strong increase in the formalization of the economy.

Taxation in Ecuador can be broadly categorized into personal income taxes (PIT), a value-added tax (VAT) of 12 % (food and some other goods are exempt)[3], corporate taxation (22% of profits since 2013), and a tax on foreign money transfers and special consumption taxes. Figure 2.1a gives a clear picture of the growth of tax revenue in Ecuador in the past years.[4] Between 2006 and 2015, central government tax revenues have increased from about 10% to almost 14% of GDP and have more than doubled in real terms. One of the main reasons for higher tax revenue is an increase in formalization of the economy and the tax administration's wide-ranging efforts to increase tax compliance.

[3] Following a large destructive earth quake in 2016 the Ecuadorian government increased the VAT to 14 % for the duration of one year starting in June 2016.

[4] The Ecuadorian economy was completely dollarized in 2000 following extreme hyperinflation.



(a) Tax Revenue to GDP

(b) Number of Tax Declarations

Figure 2.1: Formalization

The government has adopted a number of policies to increase formalization of the economy, the most important of which are extensive deduction possibilities of income tax. Along with 'receipt lotteries', in which citizens have the possibility to submit receipts and win prizes, these policies substantially increase the demand for receipts. Emitting receipts is not only linked to paying more VAT but also to taking part in other aspects of the formal economy such as retaining income tax and social security contributions for employees. The receipts handed in to the authorities are used to cross-check the sales of businesses and fight tax fraud, especially with respect to VAT reporting behavior. Further measures to increase tax compliance include improved information sharing between government agencies.

The general hike in tax revenue in Ecuador is also reflected in a strong increase in the number of taxpayers subject to personal income taxation. Figure 2.1b gives an overview of the absolute number of tax declarations submitted. Between 2006 and 2015, the total number of tax declarations for private sector employees increased from 1 Million to about 2.5 Million.

2.3.1 Personal Income Taxes

Ecuador has a unified PIT schedule which is levied on almost all regular sources of wage and self-employed income.[5] Tax liability in Ecuador is individually determined (no family taxation).[6]

The PIT liability is calculated progressively with numerous small jumps in the marginal tax

[5] Notable exceptions include all forms of payments from the social security system (pension payments, educational stipends, disability benefits, etc.), severance payments, interest on savings accounts, occasional capital gains, returns from investment funds or long-term deposits as well as certain additional wage benefits mandatory under labor market regulations.

[6] Furthermore, employees in the private sector pay 9.35% of their wage income in social security contributions. Paying these social security contributions entitles people to a range of benefits including pensions, health insurance, disability insurance and unemployment benefits. Social security contributions are only levied on regular wage income, not irregular special payments such as boni. Since 2014, the contribution has increased to is 9.45%. The employer pays a slightly larger share of 11.15%, constant over time.

rate, starting at 5% and going up to 35%. In 2008, the government enacted a series of reforms of the tax system, including an increase of the maximum marginal tax rate from 25% to 35%. Figure 2.2 gives an overview of the marginal tax rates in 2013. The cutoff income levels change yearly according to inflation[7], the exact values since 2006 are displayed in Table 2.1.

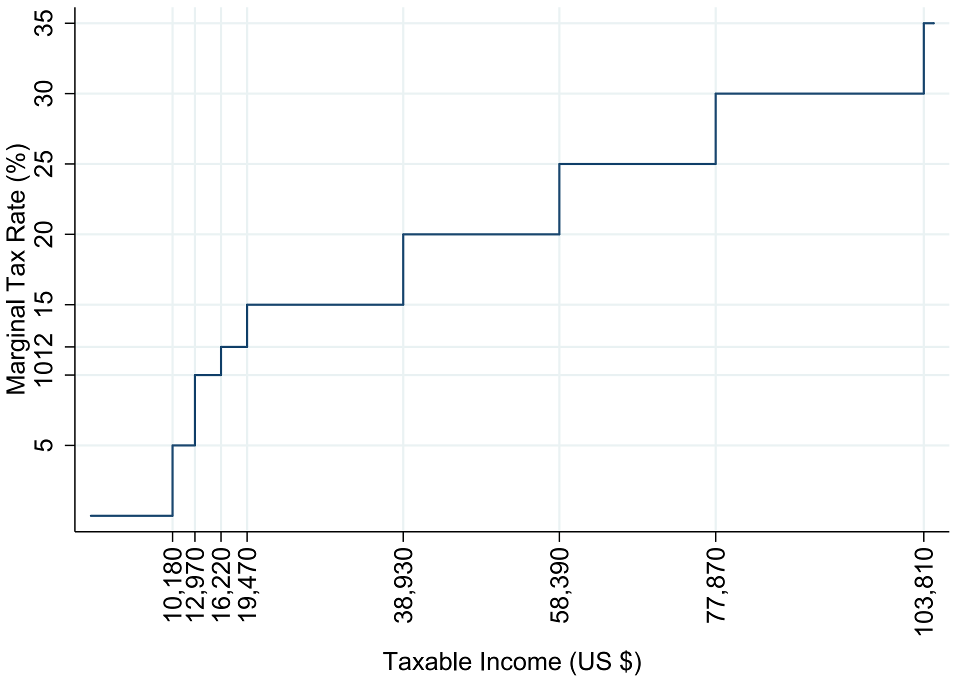


Figure 2.2: Marginal Tax Rates 2013

PIT in Ecuador starts being levied at relatively high levels. In 2013, annual income below 10,180 USD was not charged any income tax. For the same year, the monthly minimum wage is set at 318 USD, corresponding to yearly taxable income of 3,816 USD, well below the first tax bracket. The minimum wage is estimated to be slightly above the median wage and slightly below average wage in Ecuador for 2008 to 2012 (Canelas, 2014). This shows that PIT is only applicable to relatively high-earning individuals in Ecuador.

A uniqueness of the Ecuadorian tax system are the generous deduction possibilities for personal expenses in education, health, food, clothing and housing introduced in 2008. The total deductible amount of personal expenses is limited to the smaller of 50% of individual income or 1.3 times the tax-exempt income amount (in 2013 this was $1.3 \times 10{,}180 = 13{,}234$ USD). Each category is individually capped at 0.325 times the tax-exempt income amount, except for health expenditures, which have an upper limit of 1.3 times the tax-exempt amount. To make receipts presentable to the tax authority, they must be issued to the name of the tax payer or his/her dependents and include their unique identification number. One main

[7] The rate used for inflation adjustments is the yearly change in consumer price index for urban areas published by Ecuador's National Statistics Institute INEC on November 30 of a given year.

Table 2.1: Tax Brackets (in US $)

Marginal Rate	06	07	08	09	10	11	12	13	14	15
5%	7,680	7,850	7,850	8,570	8,910	9,210	9,720	10,180	10,410	10,800
10%	15,360	15,700	10,000	10,910	11,350	11,730	12,380	12,970	13,270	13,770
12%	–	–	12,500	13,640	14,190	14,670	15,480	16,220	16,590	17,210
15%	30,720	31,400	15,000	16,370	17,030	17,610	18,580	19,470	19,920	20,670
20%	46,080	47,100	30,000	32,740	34,060	35,210	37,160	38,930	39,830	41,330
25%	61,440	62,800	45,000	49,110	51,080	52,810	55,730	58,390	59,730	61,980
30%	–	–	60,000	65,480	68,110	70,420	74,320	77,870	79,660	82,660
35%	–	–	80,000	87,300	90,810	93,890	99,080	103,810	106,200	110,190

Note: Columns denote the years to which the tax brackets apply. The numbers indicate the value of the lower bound above which income is taxed at the relevant marginal rate. For example: In 2014, all income between 10,410 USD and 13,270 USD is taxed at the marginal rate of 5%.

policy objective of these deduction possibilities is to increase formalization of the economy, as wage earners have an incentive to demand receipts. In order to claim these deductions, taxpayers are legally obliged to keep copies of their receipts. The standard tax declaration form F107 submitted by the firm, however, only contains information on the total yearly amount of personal expenses in each category. If the total value of deductions exceeds a certain reporting threshold, the tax authority asks the taxpayers to additionally submit an online annex with details about the receipts.[8]

The mechanism by which tax declarations and deductions are submitted in Ecuador deserves some special attention and is key to understanding the findings in our analysis. Personal income tax is primarily filed on a firm-reported form (F107, see figure 2.A3 in the Appendix). This form can only be submitted to the tax authority by the employing firm and includes the level of deductions in personal expenses. In March of each year, wage earners fill out a form with their *projected* expenses in health, education, food, clothing and housing for that whole year and submit it to their employer. Based on these figures, the employer computes the level of the withholding tax for the following year. Workers are given the opportunity to update their information on deductions in October. If an individual claims deductions above the reporting threshold (50% of the tax free amount, or 5090 US$ in 2013), he must submit the receipts with the unique receipt number via an online annex after the end of the fiscal year[9]. While the ultimate responsibility for the overall correctness of these deductions lies solely with the employee, this system induces a unique form of third-party reporting of deductions. Recent literature shows that third-party information reporting by firms is a key driver for sustaining high levels of taxation (Kleven, Kreiner, and Saez, 2016).

For the vast majority of employees (87% of our observations), taxes and personal deductions are only reported by the employer. The remaining 13% of all observations additionally submit a self-reported tax declaration (form F102). The primary purpose of this self-reported tax declaration form is to report self-employment income. However, some individuals who additionally submit a self-reported income declaration actually do not report any self-employment income.[10]

[8]From 2008 to 2010, this threshold was $7500 and since 2011 the tax authority applies the threshhold 50% of the tax-free amount (hence 5090 US$ in 2013).

[9]The fiscal year corresponds to the calendar year.

[10]In related work, we are analyzing how individuals are using these self-reported tax declaration forms to circumvent their employer and change their level of deductions. Self-employed individuals need to file the self-reported tax declaration with their total income in March of the year following the relevant fiscal year. The exact date depends on the individual identity number and lies in between March 10th and 28th. Self-employed are liable to pay personal income tax on all of their business profits and wage income and have the same deduction possibilities as wage earners. Each summer, they are charged an advance of 50% of the previous year's tax liability.

2.4 Data and Descriptives

The data we use in this paper results from the merges of several administrative datasets in Ecuador administered by the Ecuadorian tax authority *Servicio de Rentas Internas* (SRI). The core data consist of firm-reported personal income tax returns of regular employees (tax form F107) for the years 2006-2015.

We augment these tax records by two important administrative datasets. First, we use the Ecuadorian civil registry (*Registro Civil*) that provides a range of socio-demographic variables, including the year of birth, highest level of education and gender. Second, we merge the tax returns to the central firm-level registry in Ecuador (*Catastro de RUC*). This registry contains firm-level data on industry affiliation, sector (public or private), time of formation of the firm and place of registry. We end up with detailed matched employer-employee data that allows us to track a given individual's coworkers over time.

A significant fraction of workers has multiple observations per year due to the fact that people have various employers throughout a given calender year (each employer submits one declaration per employee). To compute annual earnings we sum up the incomes at different employers for each individual and year. We consider the spell with the highest earnings as the main employer. We deflate all earnings to real 2013 USD values using the consumer price index of the Ecuadorian National Statistics Institute INEC.

For our analysis of tax responses, we exclude all individuals who are employed in the public sector and only focus on private sector employees for two important reasons. First, private sector employees might have better opportunities to adjust their taxable income by bargaining with their employer about the wages and employers in the private sector might provide more support in filing the deductions. Second, public sector employees face different incentives than private sector employees and their pay is often regulated by predetermined government pay scales.

Figure 2.3 displays the distribution of gross income in Ecuador pooling all observations in our sample from 2006 to 2015. We concentrate on workers who earn at least twelve times the monthly Ecuadorian minimum wage (yearly earnings of $12 \times 318 = 3{,}816$ USD in 2013) and those who earn less than 30,000 USD. The individual data is compressed into bins of \$50 and plotted as bin frequencies for each bin. In general, the income distribution is downward sloping, with the most frequent points being around the minimum wage. The graph contrasts the income distribution with the marginal tax schedule, as given by the step function with values on the right vertical axis. The gross income distribution is clearly smooth around all kink points of the marginal tax schedule depicted in the figure.

This is different for taxable income, i.e., gross income minus all deductions, displayed in Figure 2.4. There is a clear spike in the distribution of taxable income just before the first kink in the marginal tax schedule at 10,180 USD. Evidently, individuals do not change their

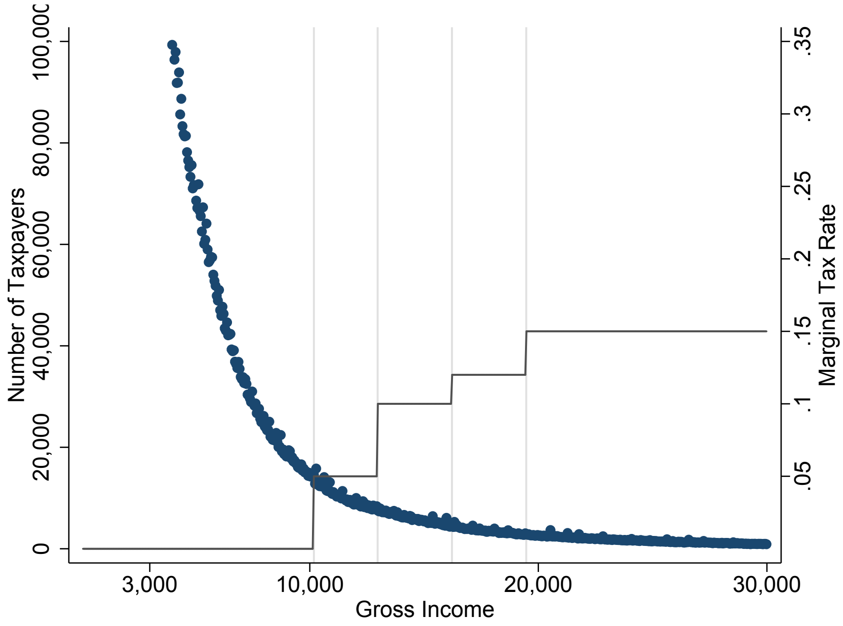


Figure 2.3: Binned Gross Income

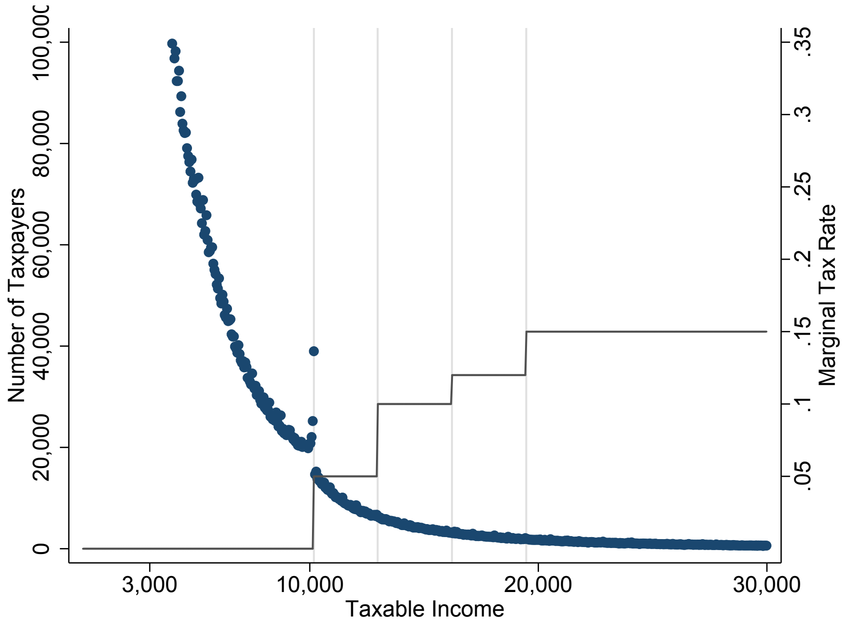


Figure 2.4: Binned Taxable Income

real labor supply but change the amount of deductions in response to the tax incentives.

While bunching is strong and pronounced at the first jump in the marginal tax schedule, we do not observe any bunching at later kink points. This could be due to the fact that the first kink where an individual starts paying taxes is the most salient. Arguably, due to behavioral biases the first dollar in taxes an individual pays can lead to higher disutility than further tax payments. Moreover, individuals may perceive a discontinuity in audit probabilities at the threshold of paying taxes and prefer to stay under the radar of the tax authority. In our analysis of bunching behavior in the following section, we therefore focus exclusively on the first kink of the marginal tax schedule.

The difference between Figures 2.3 and 2.4 indicates that adjustments in taxable income are entirely driven by reporting behavior. In particular, the introduction of the generous deduction possibilities in Ecuador in 2008 led to a wedge between the number of individuals with *gross* income above the first kink in the tax schedule and those with *taxable* income above the first kink (see Figure 2.5). Over time, a growing number of individuals avoids paying taxes by adjusting the taxable income. In the following section we quantify the amount of bunching and analyze the determinants of learning about tax avoidance.

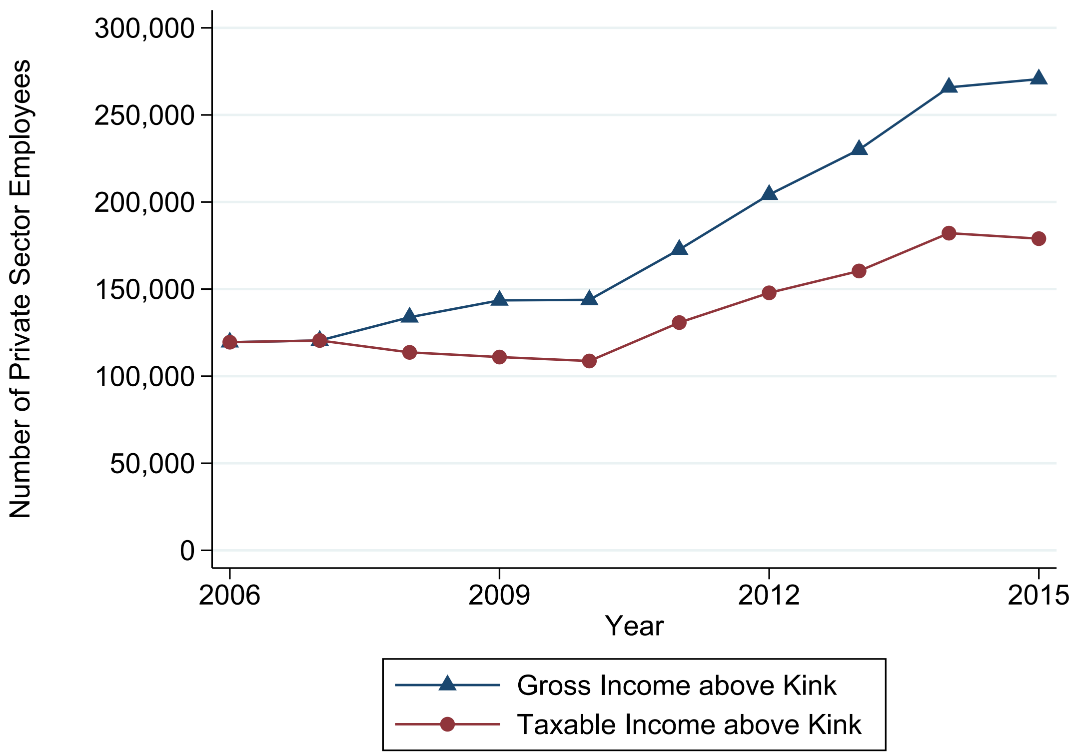


Figure 2.5: Number of Employees

2.5 Results

In this section, we present the empirical results from analyzing the individual tax return data in Ecuador. The first part uses the bunching methodology developed by Saez (2010) and Chetty et al. (2011) to estimate the extent of behavioral responses to taxation and documents general learning dynamics. The second and third part of this section analyze the channels through which this learning takes place by focusing on two main mechanisms: adapting to the firm-level practices and learning from co-workers.

2.5.1 Tax Bunching

To quantify the amount of bunching at the first kink of the marginal tax schedule, we draw on the methods laid out in Saez (2010) and Chetty et al. (2011). Using binned income data (50$ bin size), we estimate a counterfactual density (polynomial of degree 5) around the kink that would prevail in the absence of the kink and compute the difference between the actual density and the counterfactual density.[11] Figure 2.6 displays the distribution of taxable income around the kink. The empirical density is represented by the blue dots and the estimated counterfactual is represented by the red line. The estimate for the excess mass is highly significant and very large, indicating that more than three times as many individuals are located around the kink compared to the expected mass under the counterfactual of no kink.

Table 2.2 displays the estimated excess mass separately for each year in the sample period. We find positive and significant bunching in taxable income and over time the estimates of the excess mass increase strongly from 1.36 in 2006 to 6.03 in 2015. In 2006 and 2007, before the introduction of the deduction possibilities, our bunching estimates in taxable income are identical to those of gross income. Starting in 2008, however, bunching in taxable income increases strongly while we do not observe significant bunching in gross income anymore.

We employ two different strategies in order to analyze whether the overall increase in tax bunching in Ecuador is driven by experience in filing taxes. First, we show the increase in the excess mass separately for each cohort that enters the formal economy. In particular, we examine the degree of bunching for the subgroup of individuals that is observed for the first time in a particular year and follow this cohort over time. In order to hold the sample composition constant within cohorts, we only consider individuals who are observed without interruption once they entered the formal economy. Table 2.3 displays bunching estimates over time for our cohort analysis. Each row corresponds to one of the cohorts that entered between 2007 and 2014. The columns indicate how the level of bunching changes over time for these cohorts. For each cohort, there is a clear increase in the amount of bunching. Moreover, the estimates become more precise over time, indicating less heterogeneity within cohorts over

[11] Sensitivity checks varying the bin width, the parametric form of the polynomial and the bunching window left out in the estimation of the counterfactual density are available on request.

Table 2.2: Bunching Estimates over Time

	pooled	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Taxable Income	4.13***	1.36***	1.86***	2.88***	1.81***	3.34***	3.88***	4.44***	4.63***	5.18***	6.03***
	(0.24)	(0.37)	(0.36)	(0.49)	(0.61)	(0.54)	(0.58)	(0.72)	(0.91)	(0.77)	(0.61)
Gross Income	0.23	1.35***	1.85***	1.16**	-0.36	1.05	0.80	0.26	-0.36	-0.62	-0.33
	(0.29)	(0.38)	(0.36)	(0.59)	(0.81)	(0.75)	(0.75)	(0.94)	(1.04)	(0.99)	(0.79)

Note: This table reports bunching estimates for taxable and gross income by year and in the pooled sample. The estimates are based on binned income data (50$ bin size) and a counterfactual density using a polynomial of degree 5. Standard errors reported in parentheses, significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

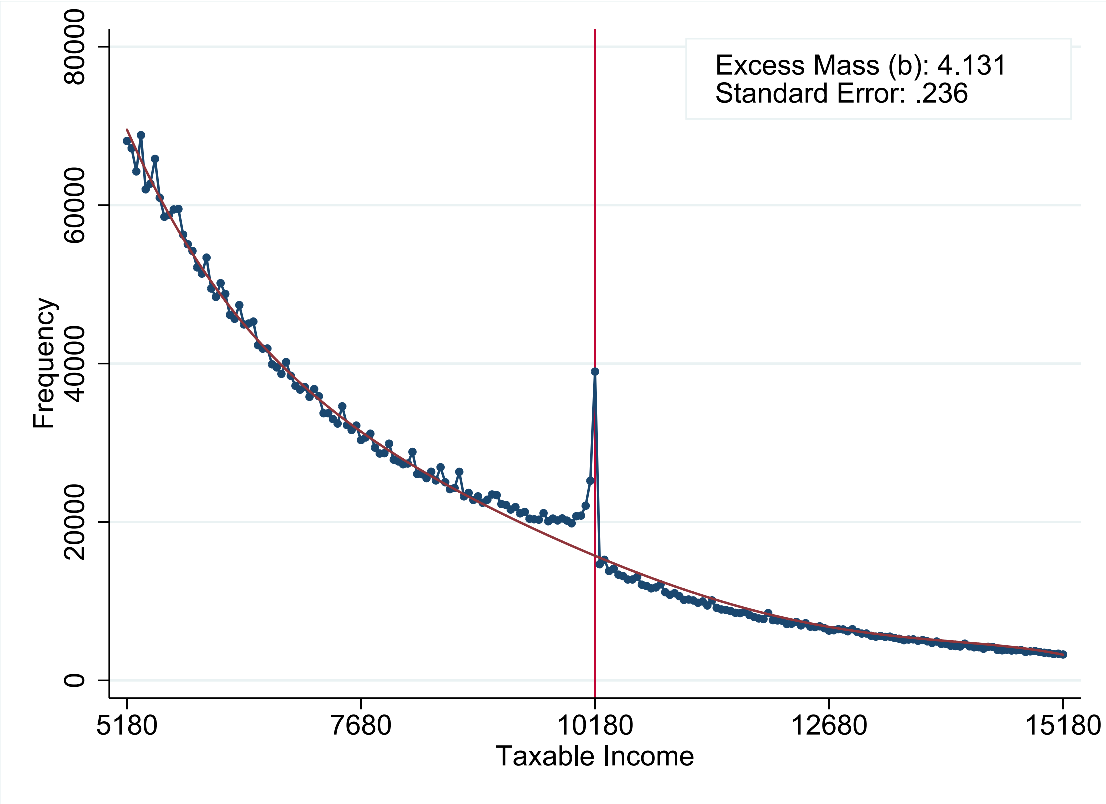


Figure 2.6: Bunching Estimates Taxable Income

years. Individuals entering the formal economy in 2007 for instance had a modest (and insignificant) excess mass of 2.59 while the same individuals had an excess mass of 6.65 in 2015. The pattern is similar for cohorts that entered the formal economy in later years. In general, bunching levels are higher in 2007 and 2008 before the introduction of deduction possibilities. Afterwards, learning did not only occur within cohorts but also across cohorts as individuals entering the formal economy in later years tend to have higher degrees of bunching.

As a second strategy to evaluate the effect of experience on the amount of bunching, we construct a specific measure of experience with the tax system. Our experience measure keeps track of whether individuals have earned more than the tax exempt threshold in the previous two years. This measure is important since only individuals who earn more than the income threshold of the first kink have an incentive to learn about deduction possibilities in order to avoid paying taxes.

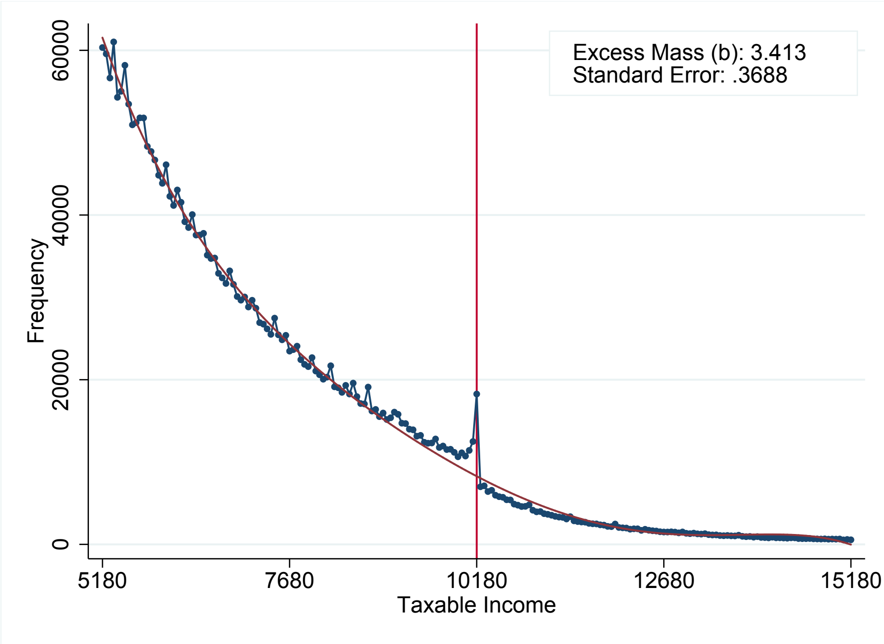
Figure 2.7 depicts and quantifies the amount of tax bunching for individuals with and without recent exposure to the tax system. In Panel (a) we observe that individuals who have not had any gross income above the first kink of the marginal tax schedule in the previous two years show rather low levels of bunching. Those individuals with at least one year of gross income above the first kink in the previous two years, however, show much stronger stronger levels of bunching. The mass in the vicinity of the kink is estimated to be 6.171 times higher

Table 2.3: Bunching Estimates over Time by Cohort

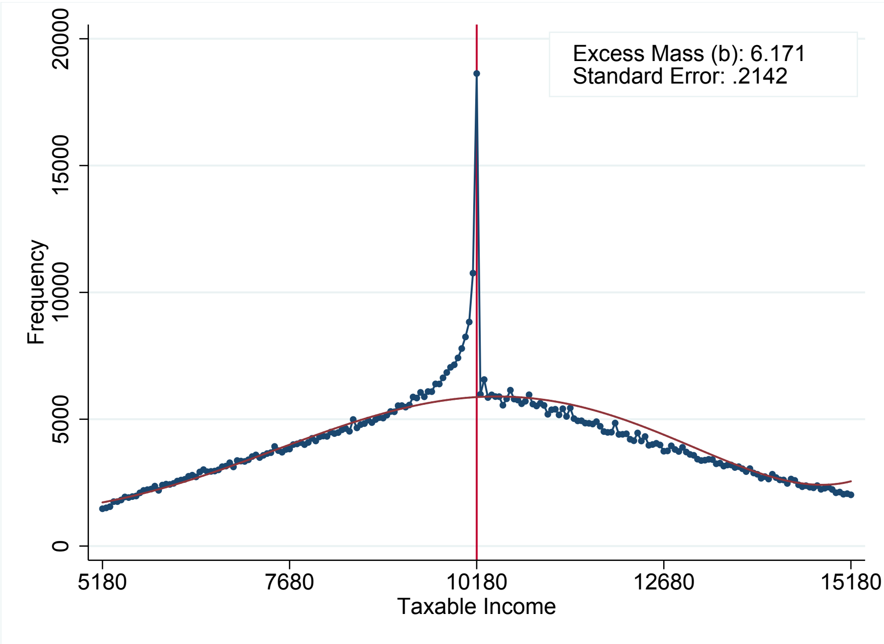
Cohort	2007	2008	2009	2010	2011	2012	2013	2014	2015	Observations
Taxable Income										
2007	2.59^{*}	2.95^{***}	2.89^{***}	3.08^{***}	4.25^{***}	4.98^{***}	4.31^{***}	4.93^{***}	6.65^{***}	48,570
	(1.50)	(1.08)	(1.08)	(0.77)	(0.74)	(0.70)	(0.58)	(0.60)	(0.65)	
2008		3.44^{**}	-0.57	2.90^{***}	2.64^{***}	4.78^{***}	3.08^{***}	4.72^{***}	3.83^{***}	79,785
		(1.59)	(0.92)	(0.75)	(0.65)	(0.68)	(0.56)	(0.51)	(0.52)	
2009			0.26	0.75	2.26^{**}	5.74^{***}	4.34^{***}	5.67^{***}	5.61^{***}	59,427
			(0.66)	(1.60)	(1.02)	(1.02)	(1.03)	(0.70)	(0.79)	
2010				0.62	2.16	3.94^{***}	4.75^{***}	5.45^{***}	5.56^{***}	67,024
				(0.98)	(1.74)	(1.21)	(1.19)	(1.00)	(0.82)	
2011					1.18	3.72^{*}	6.05^{***}	6.15^{***}	7.19^{***}	108,496
					(0.97)	(2.15)	(1.61)	(1.15)	(1.04)	
2012						2.91	4.64^{*}	5.69^{***}	5.49^{***}	140,777
						(3.23)	(2.57)	(1.35)	(0.96)	
2013							5.21	4.08^{*}	6.25^{***}	168,952
							(3.43)	(2.19)	(1.38)	
2014								3.73	7.38^{***}	219,543
								(3.07)	(1.78)	
Gross Income										
2007	2.56^{*}	1.68	1.15	1.81^{*}	1.59^{**}	0.72	0.64	0.15	0.58	48,570
	(1.50)	(1.11)	(1.14)	(0.94)	(0.86)	(0.80)	(0.77)	(0.71)	(0.79)	
2008		1.85	-2.24^{**}	1.06	0.98	1.50^{**}	0.04	0.63	-0.46	79,785
		(1.68)	(1.03)	(0.79)	(0.76)	(0.76)	(0.63)	(0.64)	(0.61)	
2009			1.25	-1.54	-0.73	2.30^{**}	0.05	0.55	-0.02	59,427
			(3.67)	(1.57)	(1.09)	(1.04)	(1.14)	(0.83)	(0.87)	
2010				1.28	-1.06	1.27	0.47	0.43	0.18	67,024
				(3.43)	(1.75)	(1.30)	(1.27)	(1.08)	(0.94)	
2011					0.20	-1.19	-0.87	-0.08	0.41	108,496
					(3.33)	(2.19)	(1.69)	(1.30)	(1.10)	
2012						-2.05	-0.78	-0.46	-1.06	140,777
						(3.28)	(2.65)	(1.52)	(1.13)	
2013							-2.57	-2.39	-0.89	168,952
							(3.36)	(2.36)	(1.35)	
2014								-3.72	-1.18	219,543
								(3.10)	(1.91)	

Note: This table reports bunching estimates for taxable and gross income by year conditioned on the cohort of entry into the formal economy. The estimates are based on binned income data (50$ bin size) and a counterfactual density using a polynomial of degree 5. Standard errors reported in parentheses, significance levels are given by $^{*} < 0.1$, $^{**} < 0.05$, and $^{***} < 0.01$.

than the counterfactual.



(a) No Gross Income above First Kink in Previous 2 Years



(b) At Least One Year of Gross Income above First Kink in Previous 2 Years

Figure 2.7: Experience in Paying Taxes

One major concern in comparing bunching estimates between these two subgroups is that they may be selected with regard to income and other socio-demographic factors.[12] To address this issue, we measure the effect of our experience measure on tax-adjustment behavior while

[12]This is partly already mitigated by the fact that the bunching estimator is a local estimator measuring the excess mass only for the specific sample at hand.

holding other factors such as income levels fixed. Table 2.4 presents results from simple probit regressions with an indicator for bunching, defined as having taxable income within the range of 1000$ to the left of the tax-exempt threshold, as the outcome variable. We restrict the sample to individuals in the years 2008 - 2015 with gross income above the kink but still within the relevant range for bunching using the deduction possibilities. Column (1) of Table 2.4 shows that our measure of experience with the tax system (defined as having earned more than the tax exempt threshold in the previous two years) has a positive and significant effect on individual bunching behavior. More importantly, column (2) illustrates that even when controlling for gross income and a range of individual and firm-level control variables, the size, direction and significance of the experience effect remains comparable. The regression furthermore provides insight into which demographic characteristics are important in determining whether a given taxpayer bunches. Woman and married individuals are more likely to bunch, and interestingly higher levels of education lead to a higher propensity to bunch.

The evidence presented in this section strongly supports the hypothesis of learning dynamics in tax bunching at the kink. Next, we turn to the question of how learning takes place and investigate firm-level responses to tax incentives.

2.5.2 Firm Dynamics

An important component of the Ecuadorian personal income tax system is that firms directly submit tax declarations to the tax authority on behalf of their employees. Moreover, even the value of deductions is jointly submitted at the workplace (see Section 2.3.1 for details). This leads to the hypothesis that firm-level practices in filing taxes have a decisive role in shaping behavioral responses to tax incentives. In this section, we provide evidence for the importance of firm-level behavior.

First, we show that the increase in bunching documented above is mainly driven by an increase in the share of firms that employ workers who use deductions in order to bunch below the first kink. We define *potential bunchers* as individuals with gross earnings in a range allowing them to lower their taxable income below the first kink of the tax schedule by using deductions. In 2013 real USD, this was gross earnings between 10180 and 20360 USD. For each firm, we calculate the share of actual bunchers among the potential bunchers as a measure of firm-level knowledge about the tax system. Analogously to the individual level cohort analysis in section 2.5.1, we then follow cohorts of firms which first appeared in the formal sector in the respective year.[13] Table 2.5 reports the extensive margin of firm bunching, i.e., the share of firms that employ at least one potential buncher who actually bunches. Evidently, there is a strong increase in the share of firms that employ bunchers over time for each of the cohorts.

[13] We restrict our sample to firms that employed potential bunchers throughout all years since their first appearance in the formal sector.

Table 2.4: Bunching Individuals

	(1)	(2)
Income Experience	0.0828***	0.0666***
	(0.0119)	(0.0136)
Gross Income		0.0000242***
		(0.00000223)
Age		0.00626***
		(0.00226)
Female		0.114***
		(0.0113)
Foreign		-0.00962
		(0.0173)
Married		0.0454***
		(0.00816)
Secondary Education		0.0346*
		(0.0197)
Tertiary Education		0.0600**
		(0.0280)
Observations	1069607	1050694

The table shows results from a probit regression with a binary indicator for bunching individuals as dependent variable. The sample is restricted to potential bunchers in 2008 to 2015. Further (unreported) control variables include age squared as well as firm-level control variables such as industry affiliation, firm-size, province, firm age and corporate firm indicator. Year fixed effects are included. Standard errors (in parentheses) are clustered at the firm level. Significance levels given by * < 0.1, ** < 0.05, and *** < 0.01.

Moreover, new cohorts start with a higher fraction of firms with bunchers than the previous cohorts had in the beginning.

Table 2.5: Extensive Margin of Firm-level Bunching over Time by Firm Cohort

Cohort	2008	2009	2010	2011	2012	2013	2014	2015	Obs
2008	0.20	0.31	0.38	0.41	0.53	0.61	0.63	0.67	489
	(0.40)	(0.46)	(0.49)	(0.49)	(0.50)	(0.49)	(0.48)	(0.47)	
2009		0.23	0.33	0.41	0.47	0.53	0.59	0.61	528
		(0.42)	(0.47)	(0.49)	(0.50)	(0.50)	(0.49)	(0.49)	
2010			0.21	0.31	0.43	0.51	0.56	0.54	555
			(0.41)	(0.46)	(0.50)	(0.50)	(0.50)	(0.50)	
2011				0.26	0.38	0.45	0.50	0.55	1100
				(0.44)	(0.49)	(0.50)	(0.50)	(0.50)	
2012					0.31	0.41	0.50	0.49	1657
					(0.46)	(0.49)	(0.50)	(0.50)	
2013						0.37	0.46	0.48	2203
						(0.48)	(0.50)	(0.50)	
2014							0.38	0.44	3280
							(0.48)	(0.50)	
2015								0.36	4847
								(0.48)	

Note: Share of firms in given cohort with at least 1 buncher. Cohorts conditioned on year of entry into formal sector and having potential bunchers in all subsequent years.

In a second step, we analyze the intensive margin of firm-level bunching and show that there is only a moderate increase in the share of bunchers within firms that started to have bunchers. Table 2.6 displays the average share of bunchers among potential bunchers in firms that have at least one buncher. In general, the share of bunchers within participating firms is relatively high (given that potential bunchers in the higher part of the income distribution would have to claim deductions at maximum values, i.e., half of their income, in order to count as a buncher). Over time, however, this share does not increase notably. In contrast, firms that enter into bunching later seem to have lower shares on the intensive margin.

In summary, the increase in overall bunching levels is primarily driven by new firms entering the set of bunching firms. Experience of the firm in the formal sector leads to a higher probability to engage into bunching on the firm level. Once a firm took the decision to allow for bunching, at least on average a relatively stable fraction of workers adjusts their income to values below the first kink. In order to gain a more detailed understanding of the mechanisms that underlie these dynamic patterns, we investigate a sample of job switchers as well as individuals with changes in their co-worker composition in the next subsections.

Table 2.6: Intensive Margin of Firm-level Bunching over Time by Firm Cohort

Cohort		2008	2009	2010	2011	2012	2013	2014	2015
2008	Share	0.62	0.51	0.49	0.48	0.44	0.46	0.44	0.49
	Std. Dev.	(0.34)	(0.32)	(0.33)	(0.33)	(0.31)	(0.31)	(0.31)	(0.31)
	Obs	96	151	187	201	258	298	310	327
2009	Share		0.59	0.53	0.53	0.52	0.51	0.50	0.50
	Std. Dev.		(0.35)	(0.34)	(0.35)	(0.33)	(0.32)	(0.34)	(0.33)
	Obs		123	174	216	247	281	312	323
2010	Share			0.68	0.55	0.57	0.55	0.52	0.53
	Std. Dev.			(0.33)	(0.32)	(0.33)	(0.33)	(0.32)	(0.32)
	Obs			115	173	238	281	312	300
2011	Share				0.70	0.63	0.61	0.61	0.61
	Std. Dev.				(0.32)	(0.33)	(0.33)	(0.34)	(0.33)
	Obs				286	417	493	553	601
2012	Share					0.75	0.68	0.66	0.66
	Std. Dev.					(0.31)	(0.32)	(0.33)	(0.34)
	Obs					508	686	822	818
2013	Share						0.81	0.73	0.73
	Std. Dev.						(0.28)	(0.32)	(0.32)
	Obs						809	1021	1049
2014	Share							0.84	0.79
	Std. Dev.							(0.27)	(0.29)
	Obs							1231	1444
2015	Share								0.88
	Std. Dev.								(0.24)
	Obs								1738

Note: Share of bunchers among potential bunchers in given cohort within firms with at least 1 buncher. Cohorts conditioned on year of entry into formal sector and having potential bunchers in all subsequent years.

2.5.3 Job Switchers

For our sample of job switchers, we consider all job transitions of individuals who switch their main employer between 2010 and 2014.[14] In the case of multiple moves of one worker in this period, we only consider the first move.[15] In order to have balanced observations for the event study outlined below, we only keep job switchers where we are able to observe at least two consecutive years before and after the move at the respective firm of origin and destination.

We classify the firm of the job switchers into quintiles based on the coworker bunching shares in their origin and destination firm. In particular, based on the sample of all private sector employees with gross earnings between 5000 and 25000 USD[16] in a given year, we compute the distribution of the share of co-workers who bunch and split the sample into quintiles. For each move, we can then assign the origin firm as well as the destination firm to one of the quintiles for the respective years.

Summary statistics for the full sample of job switchers are reported in the first column of Table 2.7. On average, an individual who changed jobs is 32 years old. 46% of the movers are married, 30% are female, and 25% of the movers have some kind of tertiary education. The average move is related to a substantial raise in wages as the mean gross income increases from about $6100 to about $6700. Similarly, taxable income increases from $5660 to $6200. The share of workers who file deductions also increases (from 8% to 10%).

Using an event study graph, we observe the dynamic adjustment process of individuals depending on the quintile they are moving to. Figure 2.8 plots the share of bunchers in taxable income, defined as those who report taxable income in a $1000 window to the left of the kink, among workers starting from a firm in the middle quintile of the bunching share distribution. The horizontal axis indicates the year relative to the move where year zero is the first year at the new firm. The data show a clear asymmetric pattern of adjustment. The share of bunchers among workers who move to a high-bunching firm sharply increases after the move with an especially strong increase in the second year at the new firm, resulting in the bunching share more than doubling its pre-move level. In contrast, the share of bunchers among workers moving to mid- or low-bunching firms both have a general upward trend in the years after the move. However, this upward trend is magnitudes smaller than the increases among individuals moving to a high bunching environment.[17]

[14] In case of multiple employers we consider the main employer as the one with the highest earnings. The year of move is the first year in which the main employer of an individual has changed.

[15] In a robustness check, we also analyze the sample of movers who move only once with no change in the results.

[16] By restricting our sample to this subset, we guarantee that we only take into account those coworkers that are close enough to the first kink for bunching to be a viable option.

[17] Table 2.A2 in the appendix depicts the same event-study graph for individuals starting in the low or high quintile of the bunching distribution. In both alternative samples we also find a much stronger increase in the share of bunchers among individuals moving to the top quintile than among individuals moving to the mid or low quintile.

Table 2.7: Job Switchers - Descriptives

	Descriptive Statistics			
	(1)	(2)	(3)	(4)
	Full Sample	Mid to Low	Mid to Mid	Mid to High
Demografics				
Age	32.29	33.27	31.27	30.75
Married	0.46	0.47	0.45	0.46
Female	0.30	0.30	0.28	0.31
Tertiary Education	0.25	0.23	0.20	0.27
Pre-Move				
Gross Income	6092.72	5868.99	6278.32	6703.97
Taxable Income	5662.10	5493.57	5838.80	6232.53
Share Deduction Filers	0.08	0.07	0.07	0.08
Buncher	0.04	0.02	0.02	0.04
Post-Move				
Gross Income	6733.50	5115.60	7037.30	7450.82
Taxable Income	6190.15	4854.24	6483.60	6748.53
Share Deduction Filers	0.10	0.06	0.09	0.14
Buncher	0.04	0.02	0.04	0.06
Observations	152617	5919	6717	5682

Notes: This table reports summary statistics for the job switcher sample, consisting of all individuals who switch their job between 2010 and 2014 (regarding only their first move) and for whom it is possible to observe at least two consecutive years before and after the move. Pre-move gives mean values in the year before the move, post-move the respective values in the first year at the new firm. Individuals are grouped into quintiles depending on their coworker bunching shares for any given year. Columns (2) to (4) represent individuals starting in the mid (third) quintile of the bunching distribution in the year before the move and moving to a firm in the low (first), mid (third) or high (fifth) quintile.

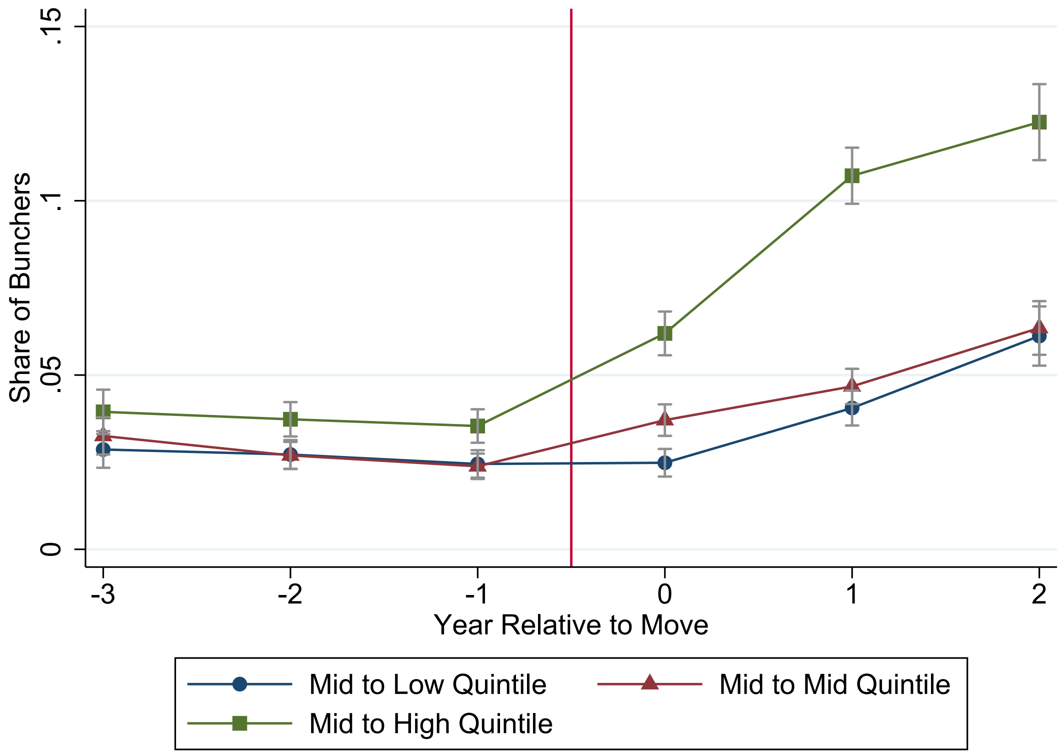


Figure 2.8: Event Study Job Switchers

Figure 2.8 indicates parallel and stable pre-move trends between individuals moving to firms in different parts of the bunching share distribution. This lends credibility to the parallel trends assumptions in standard difference-in-differences type analyses. However, columns 2-4 of Table 2.7 show that job switchers to low-, middle-, and high-bunching firms might be selected in terms of observable pre-move characteristics. In order to address possible selection issues, we employ three differing identification strategies that quantify the magnitude and significance of the effects of switching a job while controlling for individual unobserved heterogeneity as well as a number of time varying individual characteristics such as earnings before and after the job switch.

The main idea of the first identification strategy is to compare job switchers starting in a firm in the mid qunitile of the bunching distribution and moving to a firm in the high quintile to those starting in the mid quintile and moving to a firm in the same quintile. We apply the same approach to individuals moving to a firm in the low quintile of the bunching distribution. For each destination quintile $\in \{low, high\}$, we separately estimate the following regression on the subsample of individuals starting in a firm in the mid quintile and moving to the respective destination quintile:

$$Y_{it} = \beta_0 + \sum_{k=-2}^{k=2} \gamma_k D_{it}^k + \delta post_{it} \times quintile_i + \vartheta X_{it} + \lambda_t + \alpha_i + \varepsilon_{it}. \qquad (2.1)$$

The dependent variable is an indicator equal to one if individual i has taxable income within a \$1000 window to the left of the kink at time t. We include event-time dummies $D_{it}^k = 1\{t = k\}$ indicating the respective year relative to the job switch (with $k = 0$ being the first year at the new firm) in order to control for any general trends occurring in event time. The indicator variable $post_{it}$ takes on the value of one in the years after the job switch and $quintile_i$ takes on the value of one if an individual moved to a high or low quintile respectively. X_{it} are worker and firm characteristics, including gross earnings, age squared, firm size, industry classification and an indicator for corporate firm status. We further include individual (α_i) and time (λ_t) fixed effects. The coefficient δ measures the general effect of moving to a high or low bunching firm respectively. [18][19]

The estimates are displayed in Panel A of Table 2.8. Columns (1) and (3) are without and columns (2) and (4) with the individual and firm-level controls X_{it}. The results confirm very strong firm-level effects on individual tax adjustment behavior: moving to a high quintile firm increases bunching by more than 3 percentage points while moving to the low quintile has no significant effect (particularly when controlling for time-varying worker and firm characteristics).

In a second model, we explicitly look at the timing of the effects by estimating separate coefficients for each period relative to the move. Particularly, we modify the equation to

$$Y_{it} = \beta_0 + \sum_{k=-2}^{k=2} \gamma_k D_{it}^k + \sum_{k=-2}^{k=2} \delta_k D_{it}^k \times quintile_i + \vartheta X_{it} + \lambda_t + \alpha_i + \varepsilon_{it} \qquad (2.2)$$

where the coefficients δ_k on the interaction term measure the anticipatory and post treatment effects reported in Panel B of Table 2.8. Differentiating the effect by year relative to the job switch we find no anticipatory effects before the job switch throughout the samples. The effects accruing to moves to a high bunching environment are persistent and strongest in the second year after the move. In contrast, moving to a lower bunching environment has no significant effect in any year after the move.

In our third specification, we restrict the sample to those individuals who switched to a high or low bunching environment and identify the effects only through the timing of the

[18] In a sensitivity check, we estimate this same regression without individual fixed effects but instead a wide range of individual specific demographic controls (age, gender, education) and find no substantial difference in the results.

[19] We furthermore estimate the same regression without the D_{it}^k event-time indicators and find no substantial change in the direction of the results.

Table 2.8: Job Switchers

	(1)	(2)	(3)	(4)
	Mid to Low		Mid to High	
Panel A: Overall Effect				
After event year	-0.00774**	-0.00188	0.0356***	0.0314***
	(0.00386)	(0.00405)	(0.00485)	(0.00473)
Panel B				
Anticipatory Effects				
Event year - 2	0.00350	0.00332	0.00417	0.00333
	(0.00519)	(0.00519)	(0.00559)	(0.00562)
Event year - 1	0.00408	0.00525	0.00534	0.00408
	(0.00546)	(0.00542)	(0.00616)	(0.00612)
Post Treatment Effects				
Event year	-0.00906	-0.00274	0.0185**	0.0148*
	(0.00591)	(0.00597)	(0.00779)	(0.00765)
Event year + 1	-0.00288	0.00349	0.0544***	0.0488***
	(0.00666)	(0.00690)	(0.00790)	(0.00787)
Event year + 2	-0.000188	0.00561	0.0494***	0.0435***
	(0.00838)	(0.00838)	(0.0101)	(0.0100)
Observations	65224	65186	64504	64473
Panel C: Timing				
Event year - 1	-0.00272	-0.00130	-0.00212	-0.00578
	(0.00327)	(0.00544)	(0.00409)	(0.00767)
Event year	-0.00238	0.00634	0.0245***	0.0165
	(0.00337)	(0.00931)	(0.00613)	(0.0144)
Event year + 1	0.0132***	0.0212	0.0699***	0.0541**
	(0.00450)	(0.0137)	(0.00595)	(0.0231)
Controls	No	Yes	No	Yes
Observations	23560	23542	22676	22662

The panels of this table denote the results from regression equations (2.1), (2.2) and (2.3) respectively. Standard errors (in parentheses) are clustered at the destination firm by year level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

move. We do not employ a comparison group anymore. Specifically, we run the following regression:

$$Y_{it} = \beta_0 + \sum_{k=-1}^{k=2} \gamma_k D_{it}^k + \vartheta X_{it} + \lambda_t + \alpha_i + \varepsilon_{it} \tag{2.3}$$

with the variables as defined above. In order to rule out any compositional effects, we furthermore restrict the sample in this regression to only include observations from the two years before and after the move for which we have a perfectly balanced panel. Panel C of Table 2.8 presents the results of these additional regressions. We find very similar results to before and take this as further evidence for the robustness of our findings.

In summary, the asymmetry in the adjustments after moves into the different quintiles points towards workplace or firm driven knowledge effects. In particular, moving to an environment with a higher share of coworkers who bunch has a learning effect and increases the likelihood to bunch. This effect is persistent and strongest in the second year after the job switch. On the contrary, moving to an environment with a lower share of bunchers does not change an individual's behavior and thus is consistent with a memory effect.

The asymmetric response to firm-level bunching confirms the finding of knowledge effects in Chetty et al. (2013). They analyze moves of self-employed individuals between regions and find asymmetric responses that are also consistent with learning and memory. In particular, self-employed workers who move to a region with a high share of bunchers increase their bunching while there is no effect for movers to low-level regions.

In order to lend credibility to our results we have conducted a number of robustness checks and alternative specifications. Table 2.A1 in the Appendix shows the results from regression equation (2.1), however, here restrict the sample to those individuals with gross income in a range where they can use their deductions to bunch at the first kink of the marginal tax schedule. Even though the sample is smaller, we find no changes in the results. If anything, the magnitude of the effects is larger.

2.5.4 Co-worker Learning

The previous section on job switchers documents that firms seem to be a key driver for individual bunching behavior. Individuals learn about tax adjustment opportunities from their firms. However, this learning could be driven both through learning directly from the firm or learning from co-workers. In order to disentangle these two learning mechanisms from each other, we look specifically at how individuals respond to possible information flows provided by their co-workers. We do not find evidence for individuals learning about tax-adjustment opportunities through changes in the composition of their co-workers.

We quantify this co-worker learning channel by looking at individuals with recent changes to their co-worker composition. Specifically, we construct a sample of firms with incoming employees who were potential bunchers[20] due to their gross income in the year before joining the new firm. We only consider firms hiring new workers once in the years 2010-2014 and in which we can observe at least two years before (2008 and 2009) and two years after (2014 and 2015) the event. These restrictions provide a sample balanced in event time and allow us to abstract from various treatments happening sequentially.

Among these firms with incoming potential bunchers, we divide the new employees into those that reduced their taxable income to just below the first kink ("bunchers")[21] and those that did not in the year *before* joining the new firm. We use this distinction to classify firms into "treatment" (receiving bunchers) and "control" (receiving non-bunchers) groups.

Table 2.9 provides descriptive statistics for the workers in this sample of firms. Along key demographic variables (average age, share married, share female, share tertiary education) treatment and control groups are very similar. Furthermore, average firm size between the two groups (58 and 61 employees) is very similar. There are some differences in terms of wages and tax-filing behavior in the year before the arrival of new co-workers.

Table 2.9: Co-worker Learning - Descriptives

	(1) Full Sample	(2) Control	(3) Treatment
Demografics			
Avg. Age	35.87	35.88	35.81
Share Married	0.51	0.51	0.52
Share Female	0.37	0.36	0.39
Share Tertiary Education	0.32	0.32	0.33
Firmsize	58.68	58.27	61.35
Pre-Event			
Avg. Gross Income	7143.18	7018.27	7939.66
Avg. Taxable Income	6396.51	6301.83	7000.20
Share Deduction Filers	0.13	0.13	0.17
Share Taxable Income Buncher	0.06	0.05	0.08
Observations	3526	3048	478

Notes: This table shows descriptive statistics for the sample of firms used in the co-worker analysis. Control refers to firms receiving incoming potential bunchers that did not bunch and treatment refers to firms receiving incoming potential bunchers that did bunch in the year prior to joining their new firm. Pre-event refers to the year before the arrival of new co-workers.

Using a similar event study methodology as employed in Section 2.5.3, we plot average leave-

[20] We define potential bunchers as individuals with gross earnings in a range allowing them to lower their taxable income below the first kink of the tax schedule by using deductions. In 2013 real USD, this was gross earnings between 10180 and 20360 USD.

[21] We again take at an interval of 1000 USD to the left of the first kink.

out bunching levels in treatment and control firms relative to the year of the move. A given firm's leave-out bunching share disregards the new co-worker and only calculates the share of bunchers among the original co-workers. The results in Figure 2.9 suggest that, while workers in treatment firms tend to have higher bunching shares throughout the whole sample period, their tax adjustment behavior does not change substantially after the arrival of a buncher.

We conduct the same event study for subsamples in which we suspect the influence to be stronger. Figure 2.10 depicts firms which had no bunchers before the incoming worker and Figure 2.11 small firms (less than 25 employees). In both of these cases our original finding of no effect is confirmed.

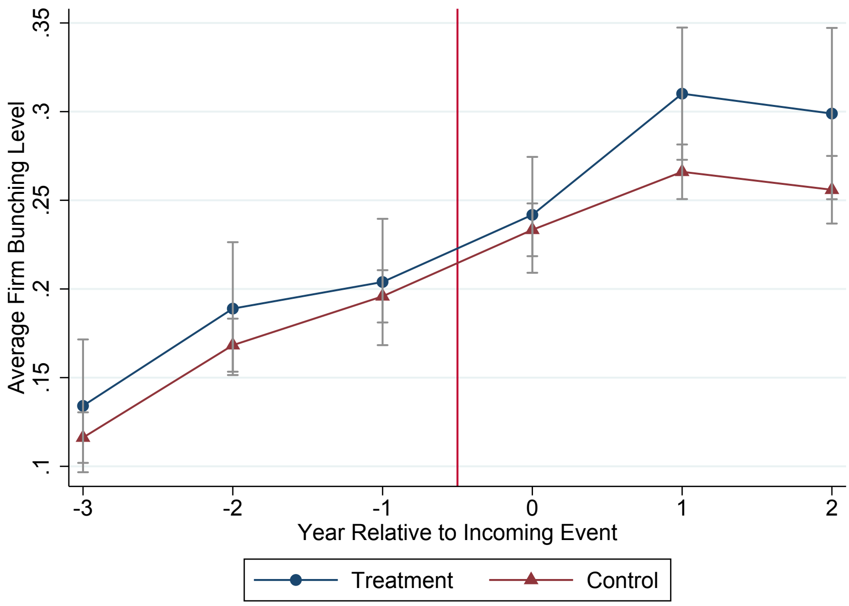


Figure 2.9: Coworker Learning - All

Table 2.10 provides regression results for the previous graphic evidence. With the aim of addressing possible selection issues and quantifying the magnitude of the effects, we mirror the identification strategies employed in Section 2.5.3. Specifically, we estimate

$$Y_{it} = \beta_0 + \sum_{k=-2}^{k=2} \gamma_k D_{it}^k + \delta post_{it} \times treat_i + \vartheta X_{it} + \lambda_t + \alpha_i + \varepsilon_{it}. \tag{2.4}$$

where i now refers to a given firm and not an individual. Y_{it} is the leave-out bunching share among the incumbent co-workers, D_{it}^k are indicators for event time, $post_{it}$ is an indicator for an observation being after the incoming co-worker, $treat_i$ is an indicator for a firm receiving an incoming buncher. We include firm (α_i) and time (λ_t) fixed effects and in X_{it} we control for firmsize as well as employee characteristics (average income, share tertiary educated, average age, share married, and share female).

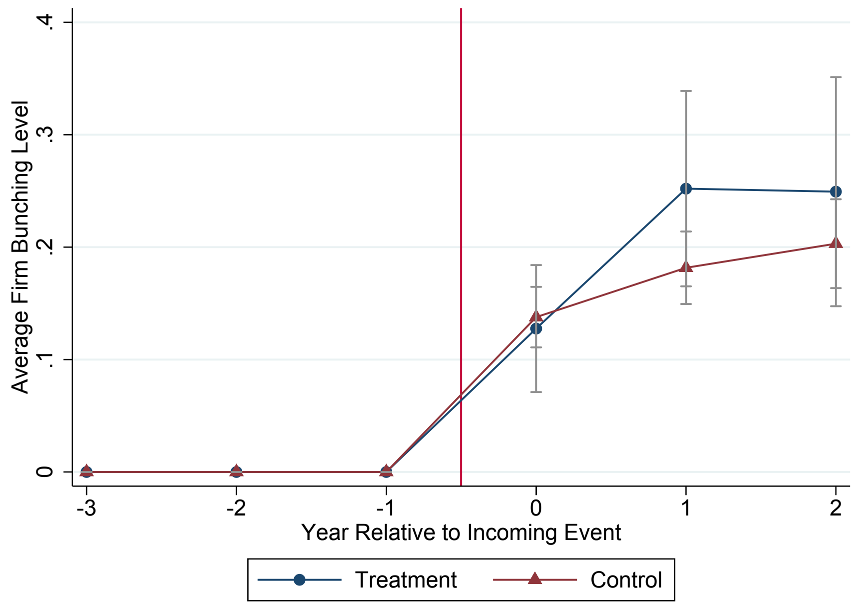


Figure 2.10: Coworker Learning - No Bunchers

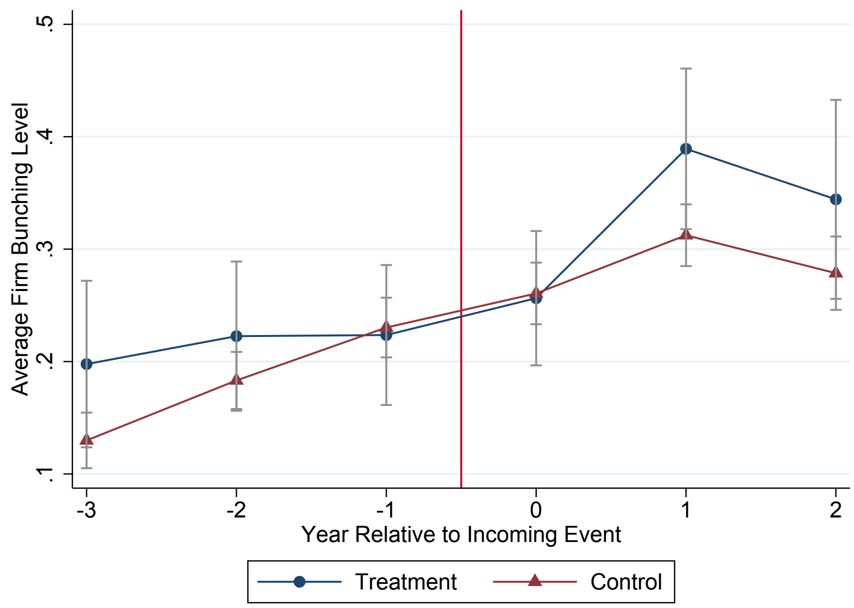


Figure 2.11: Coworker Learning - Small Firms

Table 2.10: Co-worker Learning - Regression Results

	All		No Bunchers		Small Firms	
	(1)	(2)	(3)	(4)	(5)	(6)
Overall Effect						
DiD estimate	0.0248	0.0258	0.0336	0.0252	0.0237	0.0237
	(0.0199)	(0.0199)	(0.0325)	(0.0338)	(0.0430)	(0.0431)
Anticipatory Effects						
Event year - 2	0.0137	0.0186	0.0136*	0.0100	0.0263	0.0406
	(0.0309)	(0.0311)	(0.00827)	(0.00952)	(0.0631)	(0.0640)
Event year - 1	-0.000354	0.00374	0.0136*	0.00921	-0.0184	-0.00190
	(0.0324)	(0.0326)	(0.00827)	(0.0102)	(0.0648)	(0.0657)
Post Treatment Effects						
Event year	0.00421	0.00601	0.00331	-0.00726	-0.0277	-0.0155
	(0.0308)	(0.0309)	(0.0336)	(0.0348)	(0.0648)	(0.0659)
Event year + 1	0.0480	0.0540	0.0842	0.0737	0.0731	0.0874
	(0.0334)	(0.0336)	(0.0531)	(0.0544)	(0.0731)	(0.0735)
Event year + 2	0.0447	0.0516	0.0478	0.0340	0.0406	0.0491
	(0.0391)	(0.0392)	(0.0607)	(0.0608)	(0.0855)	(0.0859)
Controls	No	Yes	No	Yes	No	Yes
Observations	11579	11574	2595	2590	4731	4731

The table reports results from regression equations (2.4) and (2.5) at the firm level. Outcome variable is the leave-out bunching share and event year refers to the year of incoming employees. Firm and year fixed effects are included throughout. We control for average income, share tertiary educated, average age, share married, share female and firmsize. Standard errors (in parentheses) are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

In a similar identification approach, we separate the overall effect into individual time components by estimating the following regression:

$$Y_{it} = \beta_0 + \sum_{k=-2}^{k=2} \gamma_k D_{it}^k + \sum_{k=-2}^{k=2} \delta_k D_{it}^k \times treat_i + \vartheta X_{it} + \lambda_t + \alpha_i + \varepsilon_{it}. \tag{2.5}$$

In this regression the coefficients δ_k measure the anticipatory and post treatment effects. These coefficients, along with the estimate for the overall effect from equation (2.4), are reported in Table 2.10. Among all three samples there do not seem to be any effects of the change in co-worker composition on individual tax-adjustment behavior.[22] We conclude these findings with the observation that learning about tax adjustment opportunities seems to be more likely driven through firm-level effects than through learning from co-workers.

The observation that firms are the main drivers of individual bunching naturally leads to the question of characterizing those firms whose employees are most likely to bunch. The following regression results in Table 2.11 show correlations between the share of bunchers in a given firm (among potential bunchers in the respective firm) and various firm-level characteristics and aggregate demographic characteristics of the employees. We see that larger firms tend to have smaller bunching shares. The sectors (industry classification) seem to play an important role in characterizing a given firm's bunching share. The reported coefficients compare a given industry with the omitted category, in this case agriculture, livestock and mining. Indeed a number of these sector coefficients go into the expected direction. Sectors with strong connections to the public sector (electricity, gas and water as well as health and social services) are related to low firm-level bunching shares.[23] The strongest positive coefficient is given by firms operating in the financial sector, as we can expect these (and their employees) to be most knowledgeable in adjusting their taxable income.

2.6 Conclusion

In this paper we analyze bunching in personal income taxes using new administrative tax-return data from Ecuador. Learning seems to play an important role in determining how individuals adjust their taxable income: people with experience and exposure to the tax system are more likely to position their taxable income within the vicinity of the first kink of the marginal tax schedule. The main margin of adjustment of taxable income lies in the reporting of generous deduction possibilities. We do not find evidence for true economic adjustments such as labor supply responses. Moreover, by exploiting data on individuals switching

[22] In unreported results we additionally identify the effect of co-workers within the sample of treated firms purely through the timing of the effect akin to the regression strategy in equation (2.3). We do not find robust evidence for any effects.

[23] Note that these results pertain only to firms in the private sector as public sector firms where excluded throughout the analysis.

Table 2.11: Bunching Firms

	Share of Bunchers in Firm	
Share Married	0.00957**	(0.00418)
Mean Age	0.00193**	(0.000857)
Share Female	0.0341***	(0.00326)
Between 10 and 100 Employees	-0.0260***	(0.00300)
Between 100 and 1000 Employees	-0.0774***	(0.00787)
More than 1000 Employees	-0.127***	(0.0103)
Corporate Firm	-0.0237***	(0.00304)
Sectors		
Manufacturing	0.0178***	(0.00166)
Electricity, gas and water	-0.00690***	(0.00179)
Construction	0.0180***	(0.00152)
Trade; Repairing	0.0187***	(0.00244)
Hotel and Restaurant	0.0117***	(0.00171)
Transport, Storage, Communication	0.00741**	(0.00241)
Financial Sector	0.0283***	(0.00253)
Real Estate, Business and Renting	0.0159***	(0.00204)
Education	0.00577**	(0.00212)
Health and Social Services	-0.0115***	(0.00220)
Other	0.00526**	(0.00217)
Observations	126540	

The table reports results from an OLS regression at the firm level with the share of bunching individuals in a firm as the dependent variable. Sample and share of bunchers constructed using only potential bunchers in 2008 to 2015. Further (unreported) control variables include average age squared, share secondary and tertiary education, share foreign workers, average number of jobs among workers and firm age. Year and province fixed effects are included. The agriculture, livestock and mining sector is the omitted category. Standard errors (in parentheses) are clustered at the industry level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

their jobs, we find strong bunching spillovers at the firm level. Someone moving from a mid-bunching environment to a high-bunching environment increases their probability to bunch by 3-5 percentage points. In contrast, for someone switching to a low bunching environment, we find almost zero effect on their probability of bunching. These asymmetric effects lead us to believe that knowledge seems to be the main driver in these spillover effects at the firm level.

Apart from establishing the importance of knowledge in individual tax adjustment behavior, we constrast a further channel of information transmission: co-worker learning. By studying firms which receintly hired employees we look at how incoming bunchers affect their co-workers' tax-filing behavior. We find no evidence for incumbent employees learning from their new co-workers' behavior, even among small firms or firms without any previous bunchers. We conclude that firms, not co-workers, seem to be the main driver of tax adjustment behavior.

From a policy perspective, these findings on how taxpayers in a low-enforcement setting learn about tax adjustment and avoidance opportunities are highly relevant. A range of developing and middle-income countries have recently undergone numerous reforms aiming towards the formalization of the economy. While designing these reforms it is important to take into account how new taxpayers react to the incentives provided by the tax system over time. Our analysis has shown that firms play an important role in how knowledge about tax adjustment opportunities is spread. In devising strategies to combat tax avoidance and increase revenue, this is an important fact to keep in mind.

In future research on behavioral responses to taxation, we think it is important to focus more strongly on dynamic aspects, especially taking into account that individuals learn over time about the incentives given by the tax system. In our analysis we show that firm-level effects play an extremely important role in determining individual tax-filing behavior. In future research, it would be of great interest to quantify the role of firms in tax filing and possibly tax avoiding behavior of individuals.

References

Pierre Bachas and Mauricio Soto. Not(ch) Your Average Tax System: Corporate Taxation Under Weak Enforcement. Technical report, UC Berkeley, 2015.

Spencer Bastani and Håkan Selin. Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics*, 109:36–49, 2014.

Timothy J Besley and Torsten Persson. Taxation and Development. *Handbook of Public Economics*, 5:51–110, 2013.

Michael Best, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan. *Journal of Political Economy*, 123(6):1311–1355, 2015.

Carla Canelas. Minimum Wage and Informality in Ecuador. Technical report, UNU-WIDER Working Paper, 2014.

Paul Carrillo, M. Shahe Emran, and Anita Rivadeneira. Do Cheaters Bunch Together? Profit Taxes, Wothholding Rates and Tax Evasion. Technical report, Working Paper, 2012.

Paul Carrillo, Dina Pomeranz, and Monica Singhal. Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement. *American Economic Journal: Applied Economics*, 9(2): 144–64, April 2017.

Raj Chetty and Emmanuel Saez. Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients. *American Economic Journal: Applied Economics*, 5(1):1–31, 2013.

Raj Chetty, John Friedman, Tore Olsen, and Luigi Pistaferri. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *Quarterly Journal of Economics*, 126(2):749–804, 2011.

Raj Chetty, John N Friedman, and Emmanuel Saez. Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review*, 103(7):2683–2721, 2013.

Henrik J Kleven and Mazhar Waseem. Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics*, 128:669–723, 2013.

Henrik J Kleven, Claus T Kreiner, and Emmanuel Saez. Why can Modern Governments Tax so much? An Agency Model of Firms as Fiscal Intermediaries. *Economica*, 83(330):219–246, April 2016.

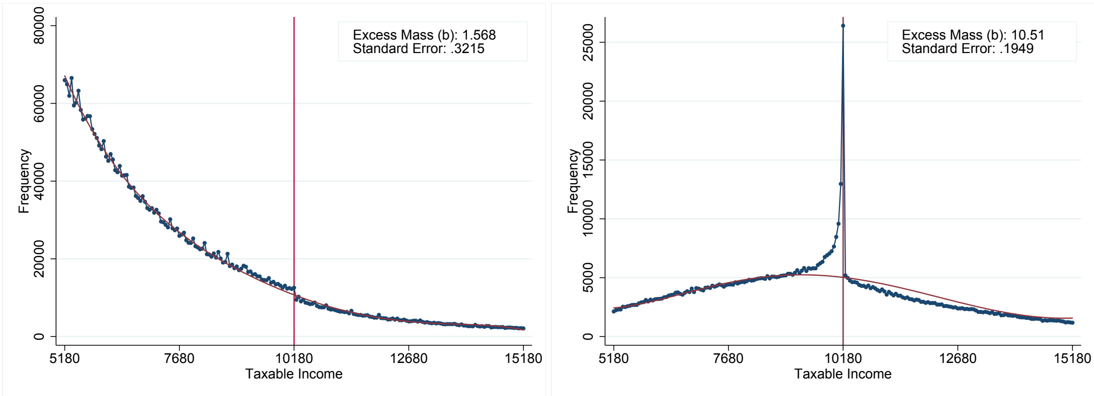
Henrik Jacobsen Kleven. Bunching. *Annual Review of Economics*, 8(1):435–464, 2016.

Jörg Paetzold and Hannes Winner. Taking the High Road? Compliance with commuter tax allowances and the role of evasion spillovers. *Journal of Public Economics*, 143:1–14, 2016.

Emmanuel Saez. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, pages 180–212, 2010.

2.A Additional Tables and Figures

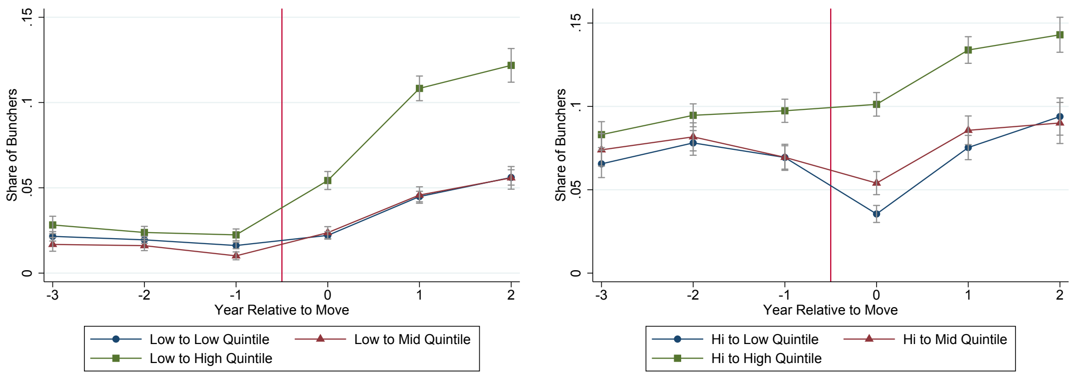
Further evidence for the fact that bunching is driven by reporting behavior can be found in Figure 2.A1. Individuals who do not file deductions for personal expenses do not display high levels of bunching (Figure 2.A1a). In contrast, individuals who file deductions (Figure 2.A1b) form a substantial excess mass to the left of the first kink in the tax schedule. The estimate here is extremely high (ten times as many individuals) and significant. Moreover, when only looking at gross income pooled in our sample period, our estimate of the bunching estimator is extremely small and insignificant (Figure 2.A5). Summing up, we find that in line with the large majority of research about behavioral responses to income taxation the reactions to tax incentives are mostly driven by reporting behavior rather than real labor supply responses. Furthermore, deductions for personal expenses are the primary tool used to avoid taxes.



(a) No Deductions

(b) With Deductions

Figure 2.A1: The Impact of Filing Deductions



(a) From Low Bunching

(b) From High Bunching

Figure 2.A2: Event Study Job Switchers

The asymmetry of the response is further emphasized by the evidence in Figure 2.A2. The left panel shows bunching shares among workers who start from a firm in the lower quintile of

Table 2.A1: Job Switchers Potential Buncher

	Mid to Low		Mid to High	
	(1)	(2)	(3)	(4)
Overall Effect				
After event year	0.00360	0.0130	0.0622***	0.0637***
	(0.0200)	(0.0221)	(0.0220)	(0.0229)
Anticipatory Effects				
Event year - 2	-0.0110	-0.0338	0.0543*	0.0404
	(0.0323)	(0.0336)	(0.0293)	(0.0310)
Event year - 1	-0.0280	-0.0423	0.0610*	0.0535
	(0.0351)	(0.0363)	(0.0354)	(0.0351)
Post Treatment Effects				
Event year	-0.0197	-0.0283	0.102***	0.0993**
	(0.0331)	(0.0349)	(0.0329)	(0.0388)
Event year + 1	-0.0198	-0.0152	0.106***	0.100***
	(0.0408)	(0.0425)	(0.0337)	(0.0348)
Event year + 2	0.0242	0.00290	0.126***	0.109**
	(0.0497)	(0.0508)	(0.0445)	(0.0460)
Controls	No	Yes	No	Yes
Observations	5493	5493	5701	5701

This table reports results for a reduced version of the job-switcher sample from Table 2.8. The sample is restricted to individuals with gross earnings between 10180 and 20360 USD, that is individuals who can use their deductions to reduce their annual income below the threshold for paying taxes. We report results from the event study-type regressions. Due to the lower number of observations we use terciles instead of quintiles. The regressions are run for individuals starting in the mid-tercile of the bunching distribution and moving to the low or high tercile respectively. The outcome variable is an indicator for having taxable income in an interval of 1000$ below the first kink. Standard errors (in parentheses) are clustered at the destination firm by year level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

the bunching distribution while the right panel refers to movers who start in the upper quintile. Among workers starting in the lower bunching quintile we see very similar patterns as before: individuals who move to the high quintile experience strong and sustained increases in bunching, whereas individuals moving to the low or mid quintile exhibit much smaller increases. Considering workers starting in the high bunching quintile we see some small additional increases among those going back to the high quintile, whereas taxpayers moving to the mid or low quintile have a temporary decrease in their probability to adjust their taxable income.

COMPROBANTE DE RETENCIONES EN LA FUENTE DEL IMPUESTO A LA RENTA POR INGRESOS DEL TRABAJO EN RELACIÓN DE DEPENDENCIA No.

FORMULARIO 107
RESOLUCIÓN No. NAC-DGERCGC12-00829

EJERCICIO FISCAL	102		FECHA DE ENTREGA	103	AÑO	MES	DIA

100 Identificación del Empleador (Agente de Retención)

105	RUC 0 0 1	106	RAZÓN SOCIAL O APELLIDOS Y NOMBRES COMPLETOS

200 Identificación del Trabajador (Contribuyente)

201	CÉDULA O PASAPORTE	202	APELLIDOS Y NOMBRES COMPLETOS

Liquidación del Impuesto

SUELDOS Y SALARIOS	301	+	
SOBRESUELDOS, COMISIONES, BONOS Y OTROS INGRESOS GRAVADOS	303	+	
PARTICIPACIÓN UTILIDADES	305	+	
INGRESOS GRAVADOS GENERADOS CON OTROS EMPLEADORES	307	+	
DÉCIMO TERCER SUELDO	311		
DÉCIMO CUARTO SUELDO	313		
FONDO DE RESERVA	315		
OTROS INGRESOS EN RELACIÓN DE DEPENDENCIA QUE NO CONSTITUYEN RENTA GRAVADA	317		
(-) APORTE PERSONAL IESS CON ESTE EMPLEADOR (únicamente pagado por el trabajador)	351	-	
(-) APORTE PERSONAL IESS CON OTROS EMPLEADORES (únicamente pagado por el trabajador)	353	-	
(-) DEDUCCIÓN GASTOS PERSONALES - VIVIENDA	361	-	
(-) DEDUCCIÓN GASTOS PERSONALES - SALUD	363	-	
(-) DEDUCCIÓN GASTOS PERSONALES - EDUCACIÓN	365	-	
(-) DEDUCCIÓN GASTOS PERSONALES - ALIMENTACIÓN	367	-	
(-) DEDUCCIÓN GASTOS PERSONALES - VESTIMENTA	369	-	
(-) EXONERACIÓN POR DISCAPACIDAD	371	-	
(-) EXONERACIÓN POR TERCERA EDAD	373	-	
IMPUESTO A LA RENTA ASUMIDO POR ESTE EMPLEADOR	381	+	
BASE IMPONIBLE GRAVADA **301+303+305+307-351-353-361-363-365-367-369-371-373+381 ≥ 0**	399	=	
IMPUESTO A LA RENTA CAUSADO	401	=	
VALOR DEL IMPUESTO RETENIDO Y ASUMIDO POR OTROS EMPLEADORES DURANTE EL PERÍODO DECLARADO	403		
VALOR DEL IMPUESTO ASUMIDO POR ESTE EMPLEADOR	405		
VALOR DEL IMPUESTO RETENIDO AL TRABAJADOR POR ESTE EMPLEADOR	407		
INGRESOS GRAVADOS CON ESTE EMPLEADOR (informativo) 301+303+305+381	349	=	

IMPORTANTE: Sírvase leer cada una de las siguientes instrucciones.

1.- El trabajador que, en el mismo período fiscal haya reiniciado su actividad con otro empleador, estará en la obligación de entregar el formulario 107 entregado por su anterior empleador a su nuevo empleador, para que aquel, efectúe el cálculo de las retenciones a realizarse en lo que resta del año.

2.- El campo 307 deberá ser llenado con la información registrada en el campo 349 del Formulario 107 entregado por el anterior empleador, y/o con la proyección de ingresos de otros empleadores actuales, en caso de que el empleador que registra y entrega el presente formulario haya efectuado la retención por los ingresos percibidos con éstos últimos.

3.- La deducción total por gastos personales no deberá superar el 50% del total de ingresos gravados, y en ningún caso será mayor al equivalente a 1.3 veces la fracción básica exenta de Impuesto a la Renta de personas naturales.

4.- A partir del año 2011 debe considerarse como cuantía máxima para cada tipo de gasto, el monto equivalente a la fracción básica exenta de Impuesto a la Renta en:
vivienda 0.325 veces, educación 0.325 veces, alimentación 0.325 veces, vestimenta 0.325, salud 1.3 veces.

5.- El trabajador deberá presentar el Anexo de Gastos Personales que deduzca, de cumplir las condiciones establecidas por el Servicio de Rentas Internas.

6.- De conformidad con la Resolución No. NAC-DGER2008-0566 publicada en el Registro Oficial No. 342 el 21 de mayo del 2008, el beneficio de la exoneración por tercera edad se configura a partir del ejercicio en el cual el beneficiario cumpla los 65 años de edad. El monto de la exoneración será el equivalente al doble de la fracción básica exenta de Impuesto a la Renta.

7.- A partir del año 2013, conforme lo dispuesto en la Ley Orgánica de Discapacidades el monto de la exoneración por discapacidad será el equivalente al doble de la fracción básica exenta de Impuesto a la Renta.

8.- El presente formulario constituye la declaración de Impuesto a la Renta del trabajador, siempre que durante el período declarado la persona únicamente haya prestado sus servicios en relación de dependencia con el empleador que entrega este formulario, y no existan valores de gastos personales que deban ser reliquidados. En caso de pérdida de este documento el trabajador deberá solicitar una copia a su empleador.
Por el contrario, el trabajador deberá presentar obligatoriamente su declaración de Impuesto a la Renta cuando haya obtenido rentas en relación de dependencia con dos o más empleadores o haya recibido además de su remuneración ingresos de otras fuentes como por ejemplo: rendimientos financieros, arrendamientos, ingresos por el libre ejercicio profesional, u otros ingresos, los cuales en conjunto superen la fracción básica exenta de Impuesto a la Renta de personas naturales, o cuando tenga que reliquidar gastos personales con aquellos efectivamente incurridos, teniendo presente los límites referidos en las notas 3 y 4 de este documento.

DECLARO QUE LOS DATOS PROPORCIONADOS EN ESTE DOCUMENTO SON EXACTOS Y VERDADEROS, POR LO QUE ASUMO LA RESPONSABILIDAD LEGAL QUE DE ELLA SE DERIVEN (Art. 101 de la L.R.T.I.)

FIRMA DEL AGENTE DE RETENCIÓN	FIRMA DEL TRABAJADOR CONTRIBUYENTE	FIRMA DEL CONTADOR
		199 RUC CONTADOR 0 0 1

Figure 2.A3: Tax Declaration Form 107 for Dependent Employees

FORMULARIO SRI-GP	**DECLARACIÓN DE GASTOS PERSONALES A SER UTILIZADOS POR EL EMPLEADOR EN EL CASO DE INGRESOS EN RELACION DE DEPENDENCIA**

EJERCICIO FISCAL	2 0 1 5	**CIUDAD Y FECHA DE ENTREGA/RECEPCION**	CIUDAD	AÑO	MES / DIA
			QUITO		

Información / Identificación del empleado contribuyente (a ser llenado por el empleado)

101	CEDULA O PASAPORTE	102	APELLIDOS Y NOMBRES COMPLETOS

INGRESOS GRAVADOS PROYECTADOS (sin decimotercera y decimocuarta remuneración) (ver Nota 1)			
(+) TOTAL INGRESOS GRAVADOS CON ESTE EMPLEADOR (con el empleador que más ingresos perciba)	103	USD$	
(+) TOTAL INGRESOS CON OTROS EMPLEADORES (en caso de haberlos)	104	USD$	
(=) TOTAL INGRESOS PROYECTADOS	105	**USD$**	
GASTOS PROYECTADOS			
(+) GASTOS DE VIVIENDA	106	USD$	
(+) GASTOS DE EDUCACION	107	USD$	
(+) GASTOS DE SALUD	108	USD$	
(+) GASTOS DE VESTIMENTA	109	USD$	
(+) GASTOS DE ALIMENTACION	110	USD$	
(=) TOTAL GASTOS PROYECTADOS (ver Nota 2)	111	**USD$**	

NOTAS:
1.- Cuando un contribuyente trabaje con DOS O MÁS empleadores, presentará este informe al empleador con el que perciba mayores in gresos, el que efectuará la retención considerando los ingresos gravados y deducciones (aportes personales al IESS) con todos los empleadores. Una copia certificada, con la respectiva firma y sello del empleador, será presentada a los demás empleadores para que se abstengan de efectuar retenciones sobre los pagos efectuados por concepto de remuneración del trabajo en relación de dependencia.
2. La deducción total por gastos personales no podrá superar el 50% del total de sus ingresos gravados (casillero 105), y en ningún caso será mayor al equivalente a 1.3 veces la fracción básica exenta de Impuesto a la Renta de personas naturales. A partir del año 2011 debe considerarse como cuantía máxima para cada tipo de gasto, el monto equivalente a la fracción básica exenta de Impuesto a la Renta en: vivienda 0.325 veces, educación 0.325 veces, alimentación 0.325 veces, vestimenta 0.325, salud 1.3 veces.

Identificación del Agente de Retención (a ser llenado por el empleador)

112	RUC 1 7 6 0 0 1 3 2 1 0 0 0 1	113	RAZON SOCIAL, DENOMINACION O APELLIDOS Y NOMBRES COMPLETOS **SERVICIO DE RENTAS INTERNAS**

Firmas

EMPLEADOR / AGENTE DE RETENCION	EMPLEADO CONTRIBUYENTE
	FIRMA DEL SERVIDOR

Figure 2.A4: Tax Declaration Form for Filing Deductions for Personal Expenses

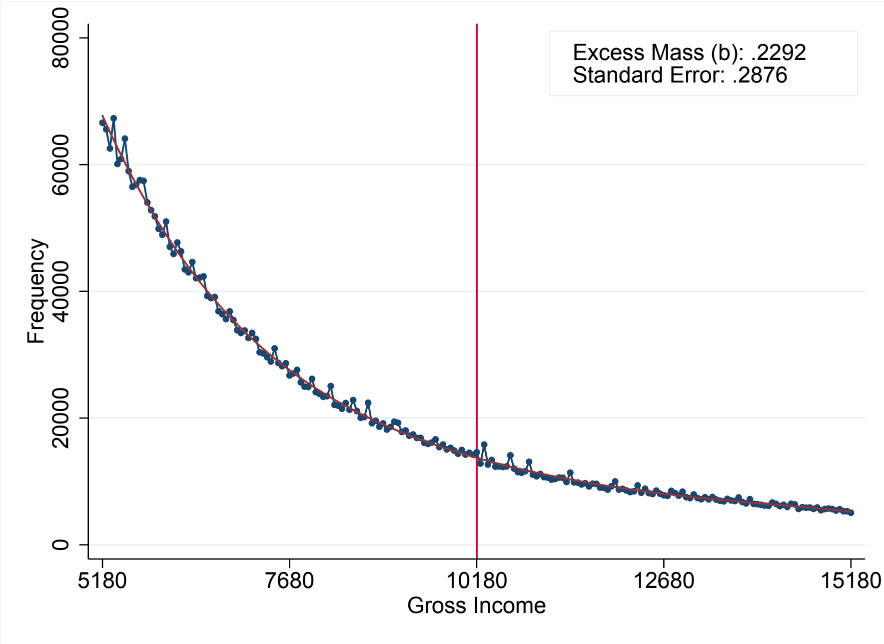


Figure 2.A5: Bunching Estimates Gross Income

Chapter 3

Sourcing, Learning, and Matching in Labor Markets

with Albrecht Glitz, Virginia Minni, and Andrea Weber

3.1 Introduction

Hiring decisions are an important determinant of firm success. Informational uncertainties about worker quality, however, are an impediment to finding good matches and hence to labor market efficiency. It is particularly complicated to screen and evaluate skills that have been acquired on the job in other firms. As a potential mechanism to reduce uncertainty around worker quality, firms can rely on information provided through various types of networks. The literature has started to explore the role of personal relations trough referrals, e.g., from former coworkers (Cingano and Rosolia, 2012; Saygin, Weber, and Weynandt, 2014; Hensvik and Nordström Skans, 2016; Glitz, 2017), from neighbors (Bayer, Ross, and Topa, 2008), from family members (Kramarz and Skans, 2014), from individuals from the same ethnic group (Dustmann, Glitz, Schönberg, and Brücker, 2016), or more theoretically in Galenianos (2013) and Galenianos (2014).

Much less emphasis, however, has been placed on the role of networks among firms directly. In particular, firms potentially gain experience from repeated interaction with each other and might be able to learn about the match quality of workers that are poached from particular firms. As an employer hires workers from a particular source firm, it learns over time their actual productivity and it discovers whether, on average, the workers hired from the source firm possess the skills the employer is looking for. The more a firm gains hiring experience, the more it should systematically hire workers from a selected group of source firms in a non-random fashion. Hence, firms decrease match uncertainty by learning where to recruit individuals with a high firm-specific human capital base taking advantage of their prior business and industry experiences. This learning mechanism can improve the matching of workers and firms by ensuring profitable and efficient job-to-job flows.

In this chapter, I document a range of empirical facts that emphasize the importance of experience of hiring firms with specific source firms. The analysis is based on matched employer-employee data from Austrian Social Security Records from which I extract the universe of job-to-job transitions in the economy. Analyzing these transitions, I first show that older firms tend to poach their workers from a narrower set of source firms. As the age of a firm increases, its hiring gets more concentrated and new hires come from a smaller number of selected firms when conditioning on firm size and growth. This result is robust when controlling for other firm characteristics, industry classes and geographical factors. Secondly, I show that a firm's acquired experience in hiring from a specific source firm leads to higher starting wages and longer tenure of workers hired from that particular firm. Having gained experience from previous interactions with the same source reduces uncertainty and therefore results in better matches. With increasing tenure, however, the information advantage disappears and workers that are hired from sources with no previous experience catch up.

To substantiate the empirical findings, I propose a theory of employer's learning where firms learn over time from which source firms to hire their workers. Workers gain labor market experience that can be relevant for alternative job opportunities and firms learn from which firms to poach their employees. The literature on learning about match quality has focused on the firm and the worker's learning about the match quality of the firm-worker pair. Instead, I separate two different learning processes. The first is the traditional learning about the idiosyncratic match quality introduced by Jovanovic (1979, 1984). The second, unexplored by the literature, is the firm's learning about which set of source firms to hire its workers from.

I develop a search and matching model with heterogeneous workers, on-the-job search and match-quality as a pure experience good. The model builds on Pries and Rogerson (2005) and Moscarini (2005) who combine variants of two benchmark models from the literature: the Jovanovic (1979, 1984) learning model and the Pissarides (1985) search and matching model. I extend the previous research in several ways to analyze the role of employer's learning about his source firms in shaping the hiring strategies of firms and the matching between jobs and workers. In the model framework, a firm learns over time about the average quality of workers from different source firms. If a firm repeatedly proves to be a good source of highly productive matches, hiring firms are more willing to employ workers from this particular source. As a result, new workers hired from the most appropriate source firms are, on average, better matched to their firm than workers hired from other firms. Hence, they earn higher starting wages and have lower separation rates. However, as workers and firms learn about their idiosyncratic match-specific productivity, low quality matches are terminated and the wage and turnover advantage dissipates over time.

The analysis contributes to several strands of the literature. First, the paper is related to the literature on firm heterogeneity in terms of age and size and matching patterns (Brown and Medoff, 1989, 2003; Brixy, Kohaut, and Schnabel, 2007; Heyman, 2007; Gabaix and Landier, 2008; Terviö, 2008; Shane, 2009; Friebel and Giannetti, 2009). It is generally found that new firms tend to have younger workers, are more likely to hire from the unemployment pool and experience greater rates of employee turnover (see, e.g., Behrends, 2007; Haltiwanger, Jarmin, and Miranda, 2013; Ouimet and Zarutskie, 2014; Janicki, Hyatt, Dinlersoz, et al., 2017). On the other hand, old and established firms are able to attract individuals that already have a job and are characterized by longer employment tenures (see, e.g., Brown and Medoff, 1989; Bergmann and Mertens, 2011; Haltiwanger, Hyatt, and McEntarfer, 2015). This empirical evidence is consistent with the theory of the employers learning over time about their source firms and improving on the matching with the new hires.

In addition, I contribute to the empirical literature on wage equations and individual earnings dynamics by adding to the debate on the effect of job experience on starting wages and wage growth and on how job experience interacts with job tenure (Altonji and Shakotko, 1987; Dustmann and Meghir, 2005; Bagger, Fontaine, Postel-Vinay, and Robin, 2014). Economic theories of on-the-job search suggest workers will move up the job ladder from lower-paying, less productive firms towards higher-paying, more productive firms (see, e.g., Burdett and Mortensen, 1998; Moscarini and Postel-Vinay, 2008, 2013, 2016). A corollary implication is that such flows will also be from young and small to old and large firms, since theory implies that firms that are more productive will be older, larger and higher paying. The implied close relationship between firm size, productivity and wages has received some mixed empirical support. I contribute to this literature by focusing on how firm's experience in hiring from a particular source firm impacts the starting wages and wage growth of the workers coming from that firm. As the firm learns about its source firms, there is less uncertainty over the success of the match and hence, there is less scope for wage growth since starting wages are already reflecting the true productivity of the match.

Finally, hiring strategies by firms and matching have also been discussed by the broader economic literature on personnel practices as in Lazear and Oyer (2012), Oyer and Schaefer (2010) and, more specifically on matching, in Andersson, Freedman, Haltiwanger, Lane, and Shaw (2009). DeVaro (2005, 2008) finds that there is a strong association between recruitment choices and starting wages where firms face a trade off between hiring speed and match quality. Moreover, the literature has considered extensively the issue of the uncertainty concerning match quality (see, e.g., Lazear, 1995; Burgess, Lane, and Stevens, 1998; Lazear and Gibbs, 2014). Lazear (1995) develops an equilibrium model where potential employees vary in terms of their riskiness, and derives predictions about which firms are good matches for risky workers. The idea is that potential employees may vary not just in their skill, the first moment of the productivity distribution, but also in the degree to which they are risky, the second moment of the productivity distribution. Within this broad literature, the novel learning process points out at the role of firms' hiring experience as an additional solution to the uncertainty problem.

The structure of the paper is as follows. In the next section, I establish empirical results concerning the role of hiring experience for wages and tenure. In Section 3.3, I set up a learning model that rationalizes the results of the empirical analysis. In Section 3.4, I calibrate the model to illustrate its empirical predictions and its congruence with the empirical facts. Section 3.5 concludes.

3.2 Empirical Findings

3.2.1 Data

The empirical analysis is based on administrative records for the universe of private sector employment in Austria. The matched employer-employee data from the Austrian Social Security Database (ASSD, see Zweimüller et al., 2009) provides detailed daily information on employment and unemployment spells as well as on other social security related states such as sickness, retirement, and maternity leave since 1972. Each individual employment spell is linked to an employer identifier and some firm-level information. Moreover, annual earnings are provided for each worker-firm combination.

I use the spell data in the ASSD to extract all job-to-job transitions that meet a range of criteria. The sample period is from 1975 to 2010. For each year, I select all individuals who start a job spell in this particular year and link the beginning job spell to the previous job spell. I only keep the transition if there are at most 30 days of non-employment between the two job spells and if both spells last for at least 30 days. I exclude transitions that are due to renamings, spin-offs, or takeovers and exclude transitions of apprentices and marginal workers using a worker flow approach (Fink et al., 2010). Finally, I only keep job transitions to firms with more than 5 employees.

I conduct two types of analysis. First, I use firm-level data to look at how firm age impacts on the number of source firms. Second, I focus on the relationship between the firm's hiring experience, on one side, and wages and spell durations, on the other side.

Firm-level Data

For the first part of the empirical analysis, I aggregate job transitions at the firm-year level. Let $i = 1, \dots, N$ denote all the firms in the economy and n_i denote the number of new hires at firm i within the annual period. Each firm i is sourcing employees from k_i other firms. Let n_{ij} denote the number of beginners at firm i that are poached from firm j and sort the firms from which firm i hires according to the number of hirings, i.e., $0 \leq n_{i1} \leq \cdots \leq n_{ik_i} \leq n_i$. I then compute the Gini coefficient of the distribution of shares as:

$$Gini_i = \begin{cases} \frac{1}{k_i}\left(k_i + 1 - 2\left(\frac{\sum_{j=1}^{k_i}(k_i+1-j)n_{ij}}{\sum_{j=1}^{k_i} n_{ij}}\right)\right) & \text{if } k_i > 1 \\ 1 & \text{if } k_i = 1 \end{cases} \tag{3.1}$$

I drop observations from the firm-level analysis where the firm hires only one worker in the respective year, i.e. $n_i = 1$.

I use two different dependent variables: the Gini coefficient, $Gini_i$, and the number of source firms, k_i. The main regressor of interest is firm age (in years) which is constructed as the difference between the current year and the first appearance of the firm in the data. In order to deal with left censoring of firm age (since data collection started only in 1972) I consider only firms that were founded after 1972. I also compute the size of the firm in the year of the move and firm growth in terms of employees between the year of the move and the last quarter of the previous year. Moreover, I collect data on industry (4-digit NACE95 code) and location (state level, NUTS-3 level). As further control variables, I compute Gini indexes for the concentration of hirings on the regional (NUTS-3) level and industry level. For instance, the regional Gini index is 0 if all hirings are from different regions and 1 if all hirings are from the same region. In the baseline version of the empirical analysis, I restrict the sample to stable firms that survive for at least 10 years and restrict the set of source firms to those firms that are likewise present during the respective time.[1]

Hiring Data

For the second part of the empirical analysis, I follow individuals who experience a job-to-job transition over the course of the spell in the new firm. I consider only hirings by firms that exist for at least ten years and from sources that exist for at least the same time as the hiring firm. Again, I restrict the sample to firms that were established after 1972. A hiring event is a job-to-job transition that satisfies the criteria specified above. For each hiring event, I compute two measures of previous hiring experience from a particular source firm as proxies for an employer's learning. The first measure is the number of hirings from the same source prior to the hiring event. The second measure is the number of years that have passed since the first hiring event from the same source. Moreover, I compute a measure of general hiring experience, namely, the average number of previous hirings from all other sources.

For each hiring event, I collect the initial daily wage in the year of the job transition and keep track of the individual's daily wages until the spell ends. The daily wage is computed as the sum of annual earnings and bonus payments (censored at the social security contribution limit) in the year divided by the number of days worked. The second outcome variable of

[1] The results, however, are not sensitive to changes in the age restriction to 15 or 20 years and to covering the full sample.

interest is an indicator that equals to one if the spell ends in the next year. Finally, I collect data on worker characteristics and employment history before the hiring event.

3.2.2 Results

Firm-level analysis

In the first part of the analysis, I examine how the set of source firms depends on the age of the hiring firm. I pool the firm-year level information over all periods. I estimate equations with the following structure by OLS:[2]

$$y_{it} = \alpha_0 + \alpha_1 f(age_{it}) + X'_{it}\beta + \gamma_i + \varepsilon_{it} \tag{3.2}$$

where y_{it} can be either $Gini_{it}$ or the number of sources, k_{it}, $f(age_{it})$ is a flexible function of firm age, and γ_i are firm-level fixed effects. Additional regressors collected in X_{it} are firm size, firm growth between t and $t-1$, the number of new hires, state level FE, NUTS-3 level FE, industry level FE and Gini coefficients that correspond to the distribution of hirings among industries and regions.[3]

Regression results for various specifications of equation (3.2) are displayed in Tables 3.1 and 3.2. There is a clear positive and significant relation between firm age and the Gini coefficient of source firm concentration (Column 1 of Table 3.1). The size of the effect could be diluted by the fact that larger firms (which are typically older) operate in various fields and therefore need workers from different sources. When controlling for unobserved firm-level heterogeneity (Column 2), firm growth (Column 3), firm size (Column 4), the number of hirings (Column 5) and the dispersion of hirings across regions and industries (Column 6) the impact of firm age becomes indeed larger. Moreover, the positive relation remains stable when considering more flexible specifications of firm age (Columns 7 and 8) and firm size (Column 9). This indicates that older firms poach their workers from a narrower set of sources, i.e. hiring gets more concentrated over a firm's lifetime. Table 3.2 reports results where the dependent variable is the number of hiring firms and displays a clear and stable negative relation to firm age.

To get a sense of the magnitude of these effects, Figures 3.1 and 3.2 show the profile of the Gini coefficient and of the number of source firms over firm age when dummies for each individual year of firm age are included in the Equation (3.2).[4] The estimated coefficients for the Gini coefficient (number of source firms) increase (decrease) almost monotonically as firm age increases. On average, ten additional years of firm age increase the Gini coefficient by 0.04 points and decrease the number of hiring firms by about one firm. These trends slightly

[2] Firms which, at any point in time, exceed 2 times the 99 percentile in firm size are excluded from the analysis.

[3] Note that I do not include year fixed effects in any specification since otherwise age and the year dummies would be collinear in the specification with year and firm FE.

[4] The omitted category is firm age equal to 1, the first full calender year of observing each firm.

Table 3.1: Firm Age and Gini Coefficient

	(1) Gini	(2) Gini	(3) Gini	(4) Gini	(5) Gini	(6) Gini	(7) Gini	(8) Gini	(9) Gini
Firm age	0.00161***	0.00173***	0.00133***	0.00239***	0.00552***	0.00342***	0.00476***	0.00574***	0.00295***
	(0.0000806)	(0.0000778)	(0.0000905)	(0.000114)	(0.000117)	(0.0000826)	(0.000190)	(0.000404)	(0.0000770)
Firm growth			0.0228***	0.0274***	0.00873***	0.00322***	0.00356***	0.00364***	0.00185***
			(0.000632)	(0.000661)	(0.000603)	(0.000450)	(0.000452)	(0.000452)	(0.000447)
Firm size				-0.0353***	-0.0683***	-0.0360***	-0.0380***	-0.0384***	
				(0.00199)	(0.00192)	(0.00130)	(0.00132)	(0.00134)	
No. of beginners					0.101***	0.0539***	0.0543***	0.0544***	0.0516***
					(0.00126)	(0.00108)	(0.00108)	(0.00108)	(0.00107)
Gini NUTS-3 reg.						0.0808***	0.0808***	0.0808***	0.0813***
						(0.00118)	(0.00118)	(0.00118)	(0.00118)
Gini 4-digit ind.						0.452***	0.452***	0.452***	0.453***
						(0.00386)	(0.00386)	(0.00386)	(0.00386)
Firm age^2							-0.0000413***	-0.000114***	
							(0.00000556)	(0.0000275)	
Firm age^3								0.00000142***	
								(0.000000541)	
Size>15									-0.0246***
									(0.00213)
Size>50									-0.0460***
									(0.00263)
Size>100									-0.0649***
									(0.00307)
Size>250									-0.0858***
									(0.00384)
Size>500									-0.0933***
									(0.00478)
Size>1000									-0.111***
									(0.00852)
Constant	0.0784***								
	(0.000871)								
Firm FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-3 region FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	354875	329323	190271	190271	190271	190271	190271	190271	190271
R^2	0.003	0.337	0.387	0.389	0.426	0.674	0.674	0.674	0.673

Note: Firms which, at any point in time, exceed 2 times the 99 percentile in firm size are excluded from the analysis. Only firms that exist for at least 10 years are included. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

Table 3.2: Firm Age and Number of Source Firms

	(1) no. source firms	(2) no. source firms	(3) no. source firms	(4) no. source firms	(5) no. source firms	(6) no. source firms	(7) no. source firms	(8) no. source firms	(9) no. source firms
Firm age	-0.0315***	-0.0940***	-0.1000***	-0.139***	-0.0525***	-0.0862***	-0.107***	-0.124***	-0.0777***
	(0.000953)	(0.00147)	(0.00172)	(0.00215)	(0.00139)	(0.00248)	(0.00512)	(0.00938)	(0.00220)
Firm growth			0.943***	0.704***	0.125***	0.0706***	0.0653***	0.0638***	0.0998***
			(0.0141)	(0.0125)	(0.00910)	(0.0120)	(0.0120)	(0.0120)	(0.0117)
Firm size				1.582***	0.652***	0.966***	0.997***	1.005***	
				(0.0392)	(0.0310)	(0.0471)	(0.0474)	(0.0476)	
No. of beginners					2.811***	4.360***	4.354***	4.352***	4.393***
					(0.0205)	(0.0428)	(0.0428)	(0.0429)	(0.0430)
Gini NUTS-3 reg.						0.0435***	0.0446***	0.0446***	0.0261
						(0.0168)	(0.0168)	(0.0168)	(0.0165)
Gini 4-digit ind.						-1.010***	-1.015***	-1.015***	-1.054***
						(0.0241)	(0.0241)	(0.0241)	(0.0240)
Firm age^2							0.000642***	0.00189***	
							(0.000146)	(0.000585)	
Firm age^3								-0.0000245**	
								(0.0000108)	
Size > 15									0.256***
									(0.0222)
Size > 50									0.742***
									(0.0409)
Size > 100									1.439***
									(0.0646)
Size > 250									2.495***
									(0.132)
Size > 500									3.551***
									(0.281)
Size > 1000									6.024***
									(0.967)
Constant	2.866***								
	(0.0178)								
Firm FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-3 region FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	742984	717434	370631	370631	370631	190271	190271	190271	190271
R^2	0.005	0.586	0.693	0.705	0.807	0.813	0.813	0.813	0.813

Note: Firms which, at any point in time, exceed 2 times the 99 percentile in firm size are excluded from the analysis. Only firms that exist for at least 10 years are included. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

level off after 30 years. Given that the Austrian economy is mainly dominated by small and medium sized firms that hire rather occasionally, this effect is relatively sizeable. Notably, the average number of source firms in the sample is 4.83 while the median firm has only three sources (see Table 3.A1 in the appendix).

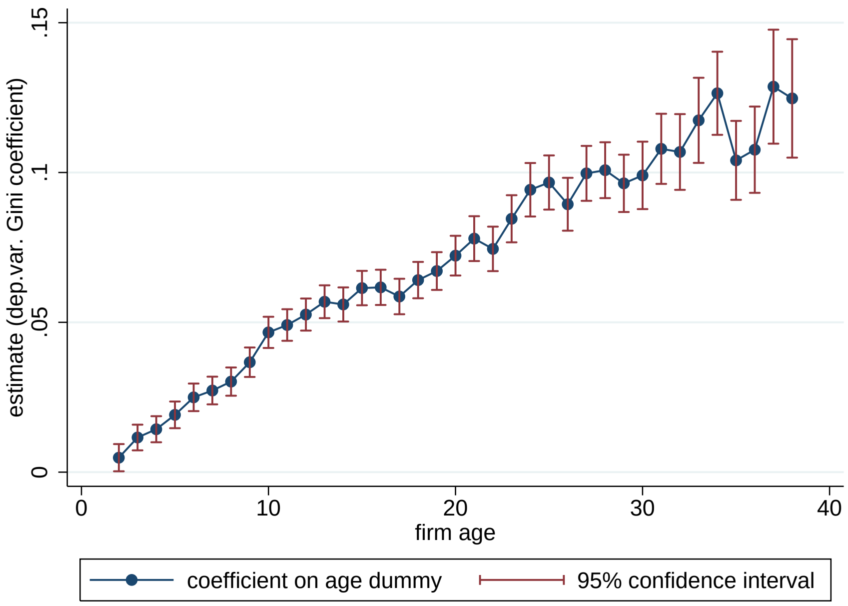


Figure 3.1: Gini Index and Age

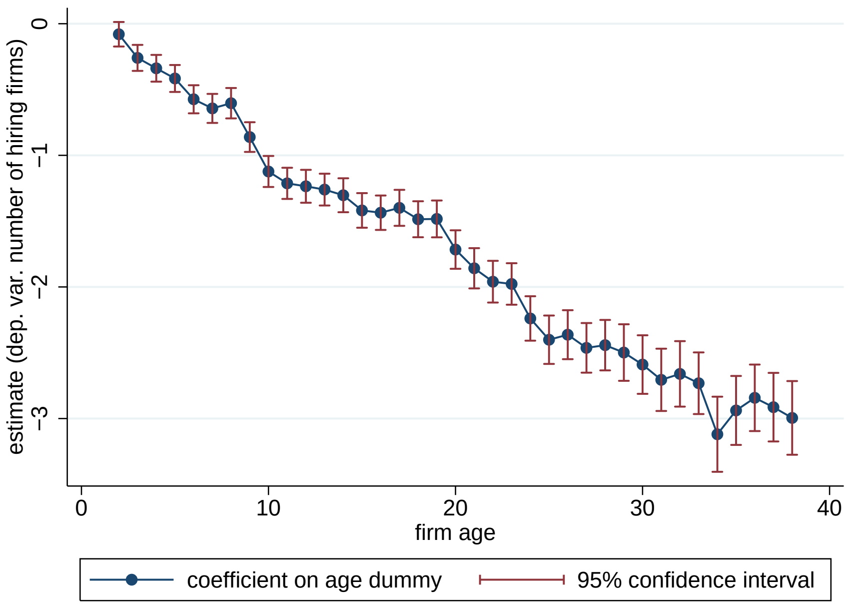


Figure 3.2: Number of Source Firms and Age

In Appendix B, I report additional regression tables and figures of various sensitivity checks.

Hiring-level analysis

In the second part of the empirical analysis, I examine the relation between the hiring experience a firm has gained with respect to a particular source firm and the wages and tenure of the workers hired from that firm. I consider all hiring events over the sample period from 1975 to 2010 and estimate equations with the following structure by OLS:

$$y_{\ell(ij,\tau),t} = \alpha_0 + \alpha_1 exper_{ij,\tau} + \alpha_2 \overline{exper}_{ij^{-1},\tau} + X'_{\ell,\tau}\beta + \gamma_i + \delta_t + \varepsilon_{\ell(ij,\tau),t} \tag{3.3}$$

where $y_{\ell(ij,\tau),t}$ is the (log) daily wage of worker ℓ hired by firm i from source firm j in year t, and τ is the year of the hiring event. The main regressor of interest, $exper_{ij,\tau}$, is one of the two measures of experience of firm i in hiring from source firm j at the time of the hiring event τ (see Section 3.2.1), and $\overline{exper}_{ij^{-1},\tau}$ is the average number of hirings firm i did from other source firms (j^{-1}) prior to the time of the hiring. I control for initial worker characteristics $X_{\ell,\tau}$ such as the daily wage prior to the hiring event, tenure at the old firm, worker age, gender, and nationality. Some versions of the model also include firm fixed effects, γ_i, year fixed effects, δ_t or firm $\times$ year fixed effects.

Tables 3.3 and 3.4 report results from estimating various versions of Equation (3.3). In Table 3.3, log daily wages of individual ℓ are regressed on the number of previous hirings from the same source firm. There is a positive and significant relation between hiring experience and wages. Of course this relation could be driven by the fact that firms with higher hiring experience are older firms that generally employ different workers than younger firms. To control for this potential confounder, I condition on the worker type by including worker characteristics and employment history in Column (2). While smaller in magnitude, the positive impact of hiring experience remains significant. Furthermore, the estimate does not change when controlling for a measure of aggregate hiring experience in Column (3). This indicates that it is indeed specific experience with a certain source firm that is important for the reduction of uncertainty and therefore the determination of wages. Finally, Columns (4) and (5) show that the positive relation between experience and wages is also present within firms, years, and firm-years by including firm and year fixed effects or firm $\times$ year fixed effects. The difference between the raw correlation in Column (1) and the more elaborate specifications indicate that part of the effect is driven by selection while a significant role of specific hiring experience is present even within firm-years. The same conclusion arises when hiring experience is measured by the number of years since the first hiring event of a worker from the same source firm (Table 3.4).

Tables 3.5 and 3.6 have the same structure as the previous tables but the outcome variable is the probability to leave the firm in the next year. This probability is negatively impacted by hiring experience, indicating that more experience leads to longer tenure, less uncertainty and better matches. While the impact of the number of previous hirings from the same source has

Table 3.3: Impact of the Number of Previous Hirings on Log Daily Wages

	(1) ln(wage)	(2) ln(wage)	(3) ln(wage)	(4) ln(wage)	(5) ln(wage)
No. prev. hirings	0.00454***	0.00198***	0.00179***	0.000674**	0.000819**
	(0.000897)	(0.000532)	(0.000544)	(0.000325)	(0.000363)
Prev. wage		0.625***	0.624***	0.409***	0.387***
		(0.00742)	(0.00731)	(0.00707)	(0.00725)
Spell duration (days)		0.0000163***	0.0000164***	0.0000164***	0.00000732***
		(0.000000986)	(0.000000968)	(0.000000825)	(0.000000745)
Age		0.0000946	0.000142	-0.00763***	-0.0318***
		(0.000304)	(0.000290)	(0.00169)	(0.000294)
Female		-0.113***	-0.114***	-0.148***	-0.130***
		(0.00503)	(0.00494)	(0.00280)	(0.00261)
Austrian		0.0728***	0.0727***	0.0422***	0.0630***
		(0.00464)	(0.00464)	(0.00214)	(0.00214)
Avg. prev. hirings			0.00121***	-0.000145	-0.0000157
			(0.000365)	(0.000132)	(0.000126)
Constant	4.066***	1.668***	1.669***		
	(0.00598)	(0.0268)	(0.0266)		
Firm FE	No	No	No	Yes	No
Year FE	No	No	No	Yes	No
Firm-year FE	No	No	No	No	Yes
Observations	5564494	5562620	5562620	5556516	5228428
R^2	0.009	0.513	0.514	0.667	0.824

Note: The wage is the log daily wage obtained as the sum of annual earnings and bonus payments (censored at the social security contribution limit) in the year divided by the number of days worked. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

Table 3.4: Impact of the Number of Years of Hiring Experience on Log Daily Wages

	(1) ln(wage)	(2) ln(wage)	(3) ln(wage)	(4) ln(wage)	(5) ln(wage)
No. prev years	0.0187***	0.00264***	0.00235***	0.00156***	0.00194***
	(0.000992)	(0.000668)	(0.000631)	(0.000274)	(0.000296)
Prev. wage		0.627***	0.625***	0.408***	0.387***
		(0.00764)	(0.00744)	(0.00721)	(0.00742)
Spell duration (days)		0.0000156***	0.0000158***	0.0000162***	0.00000718***
		(0.00000102)	(0.000000973)	(0.000000850)	(0.000000780)
Age		-0.0000998	-0.0000188	-0.00765***	-0.0321***
		(0.000350)	(0.000323)	(0.00172)	(0.000302)
Female		-0.111***	-0.112***	-0.148***	-0.130***
		(0.00545)	(0.00518)	(0.00279)	(0.00260)
Austrian		0.0740***	0.0736***	0.0423***	0.0631***
		(0.00489)	(0.00485)	(0.00215)	(0.00215)
Avg. prev. hirings			0.00180***	-0.0000839	0.0000379
			(0.000493)	(0.000117)	(0.000125)
Constant	4.051***	1.662***	1.665***		
	(0.00614)	(0.0276)	(0.0269)		
Firm FE	No	No	No	Yes	No
Year FE	No	No	No	Yes	No
Firm-year FE	No	No	No	No	Yes
Observations	5564494	5562620	5562620	5556516	5228428
R^2	0.016	0.512	0.512	0.667	0.824

Note: The wage is the log daily wage obtained as the sum of annual earnings and bonus payments (censored at the social security contribution limit) in the year divided by the number of days worked. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

no significance, this is clearly the case for the number of years since the first hiring.

Table 3.5: Impact of the Number of Hirings on the Probability to Leave the Firm

	(1) Leave	(2) Leave	(3) Leave	(4) Leave	(5) Leave
No. prev. hirings	0.0000499	-0.000125	-0.0000331	0.0000892	-0.0000304
	(0.000196)	(0.000171)	(0.000145)	(0.0000980)	(0.0000892)
Prev. wage		-0.0186***	-0.0181***	-0.0342***	-0.0291***
		(0.00274)	(0.00263)	(0.00152)	(0.00154)
Spell duration (days)		-0.0000373***	-0.0000374***	-0.0000226***	-0.0000200***
		(0.000000795)	(0.000000786)	(0.000000500)	(0.000000471)
Age		0.00384***	0.00382***	0.00914***	0.0151***
		(0.000247)	(0.000241)	(0.00124)	(0.000237)
Female		-0.0181***	-0.0179***	-0.0148***	-0.0101***
		(0.00280)	(0.00274)	(0.00141)	(0.00145)
Austrian		-0.0259***	-0.0258***	0.0253***	0.0209***
		(0.00314)	(0.00314)	(0.00147)	(0.00132)
Avg. prev. hirings			-0.000607**	0.000115	0.000230*
			(0.000244)	(0.000130)	(0.000136)
Constant	0.263***	0.377***	0.377***		
	(0.00215)	(0.0112)	(0.0109)		
Firm FE	No	No	No	Yes	No
Year FE	No	No	No	Yes	No
Firm-year FE	No	No	No	No	Yes
Observations	5642235	5640249	5640249	5634173	5298454
R^2	0.000	0.020	0.020	0.150	0.347

Note: The dependent variable is an indicator that turns on if the worker leaves the firm in period $t+1$. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

Finally, I examine how the impact of hiring experience on wages evolves over the job spell. I modify the model in Equation 3.3 to incorporate interactions between the measure of hiring experience and dummy variables for each year of tenure. In Figures 3.3 to 3.5 I display log daily wage profiles over tenure for different levels of hiring experience with respect to the source firm. All other covariates are displayed at their mean, including the firm and year fixed effects in Figure 3.4 and firm $\times$ year fixed effects in Figure 3.5. All graphs indicate a clear premium in starting wage levels for workers from source firms that are known to the hiring firm through previous hirings, conditional on worker characteristics and employment history. The increase in wages with tenure, however, is less steep for these workers such that the slope is very similar after 10 to 15 years. The course of the wage-tenure profiles is consistent with

Table 3.6: Impact of the Number of Years of Hiring Experience on the Probability to Leave the Firm

	(1) Leave	(2) Leave	(3) Leave	(4) Leave	(5) Leave
No. prev. years	0.000244	-0.00102***	-0.000925***	-0.000725***	-0.00141***
	(0.000316)	(0.000311)	(0.000305)	(0.000138)	(0.000136)
Prev. wage		-0.0185***	-0.0179***	-0.0339***	-0.0286***
		(0.00280)	(0.00264)	(0.00152)	(0.00154)
Spell duration (days)		-0.0000372***	-0.0000372***	-0.0000224***	-0.0000198***
		(0.000000805)	(0.000000787)	(0.000000499)	(0.000000471)
Age		0.00406***	0.00403***	0.00925***	0.0154***
		(0.000264)	(0.000258)	(0.00125)	(0.000228)
Female		-0.0181***	-0.0178***	-0.0148***	-0.00997***
		(0.00280)	(0.00271)	(0.00141)	(0.00146)
Austrian		-0.0261***	-0.0260***	0.0252***	0.0207***
		(0.00314)	(0.00314)	(0.00147)	(0.00132)
Avg. prev. hirings			-0.000585**	0.000133	0.000236*
			(0.000258)	(0.000132)	(0.000136)
Constant	0.263***	0.377***	0.376***		
	(0.00224)	(0.0114)	(0.0110)		
Firm FE	No	No	No	Yes	No
Year FE	No	No	No	Yes	No
Firm-year FE	No	No	No	No	Yes
Observations	5642235	5640249	5640249	5634173	5298454
R^2	0.000	0.020	0.020	0.150	0.347

Note: The dependent variable is an indicator that turns on if the worker leaves the firm in period $t+1$. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

our hypothesis of information advantages through hiring from well-known firms that are diminished by learning about individual match quality over time. In the next section, I set up a theoretical model that incorporates uncertainties about match quality and a learning mechanism about the average quality of workers from different sources in order to rationalize the empirical results.

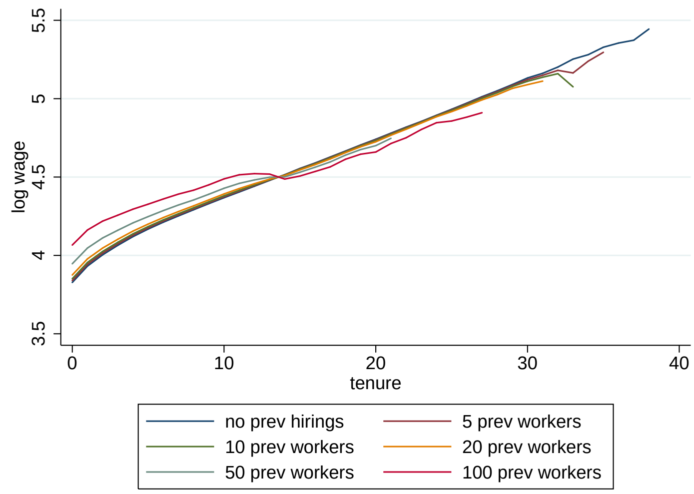


Figure 3.3: Wage-tenure Profile in Hiring Firms at Different Levels of Hiring Experience

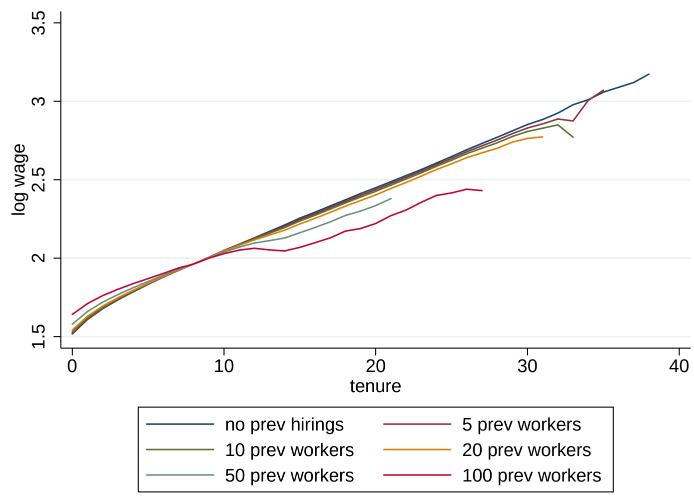


Figure 3.4: Wage-tenure Profile in Hiring Firms at Different Levels of Hiring Experience – with Firm and Year Fixed Effects

3.3 Model

This section presents the search and matching model, characterizes the equilibrium and derives its main testable predictions.

Figure 3.5: Wage-tenure Profile in Hiring Firms at Different Levels of Hiring Experience – with Firm × Year Fixed Effects

3.3.1 Environment

Time is discrete. There is a homogeneous population of risk-neutral young workers of mass n that enter the labor market with no previous labor market experience. There are three firms of three different types, $i \in \{a, b, c\}$. They discount the future at rate $\beta \in (0, 1)$. Posting a vacancy, v, to find a worker imposes a per period cost of k_v to the firm.

The workers start unemployed and are randomly employed by one of these three firms. Each firm provides the workers with some training, which is also valuable for working at one of the other firms. Hence, workers gain some labor market experience that can be relevant for alternative job opportunities. In particular, the firms are related to each another according to the network structure described in Table 3.7.

Table 3.7: Network Structure

	a	b	c
a	$0.5(\mu_l + \mu_h)$	μ_h	μ_l
b	μ_h	$0.5(\mu_l + \mu_h)$	μ_h
c	μ_l	μ_h	$0.5(\mu_l + \mu_h)$

Table 3.7 indicates that a worker trained at firm a receives relevant market experience to potentially be a high match with firm b, while the skills learned at firm a are unlikely to be valued at firm c. For simplicity, I assume that the network is symmetric. To allow for individual worker heterogeneity, I assume that match qualities are drawn from a normal distribution $N(\mu_\tau, \sigma^2)$, where $\tau \in \{l, h\}$ and $\mu_l < \mu_h$ (see Figure 3.6). Hence, training at a specific firm affects the mean of the distribution of match qualities, which changes depending on the pairs of firms. I assume that the productivity of the match between the worker and his starting firm

is fixed to be $0.5(\mu_l + \mu_h)$: the distribution of match qualities for workers who stay at their starting firm is degenerate at $0.5(\mu_l + \mu_h)$. It is important to emphasize that the different levels of the mean of the distributions, $\mu \in \{\mu_l, \mu_h\}$, are not indicative of the firms' productivities and they just represent the complementarities of the skills accumulated by the worker in one firm for production in an alternative firm.

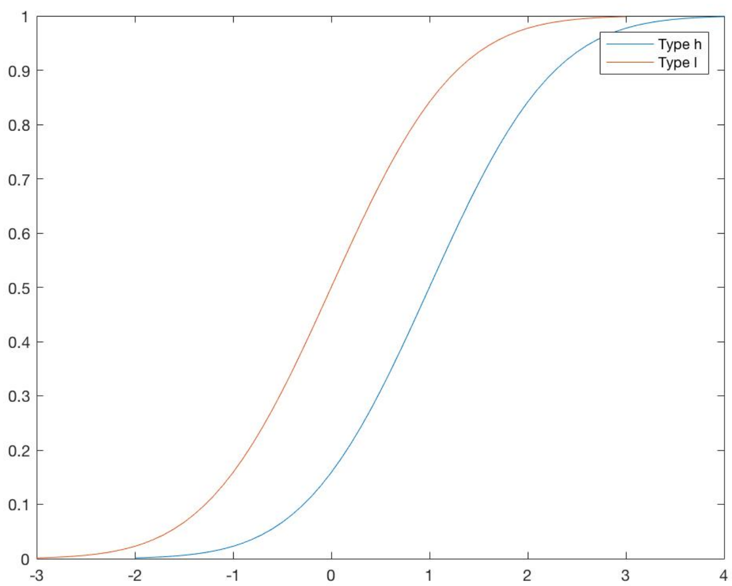


Figure 3.6: $N(\mu_\tau, \sigma^2), \tau \in \{l, h\}$

To summarize, when firm *a* meets a worker from firm *b*, the worker and the firm draw the quality of their match from the distribution $N(\mu_h, \sigma^2)$. Instead, when firm *a* meets a worker from firm *c*, the worker and the firm draw the quality of their match from the distribution $N(\mu_l, \sigma^2)$. Finally, the match quality between firm *a* and a worker trained at firm *a* is $0.5(\mu_l + \mu_h)$.

I focus on job-to-job transitions. Workers can be unemployed, in a "training state" at the starting firm or they can find a job at a different firm. I introduce two frictions. First, information frictions, whereby workers and firms do not know the worker's type. In particular, workers are ex-ante homogeneous. Once each worker gets his first job, his match quality with the other firms is unknown: the match is entirely an experience good. However, the match quality of workers that are at their starting firm is known and takes the value of $0.5(\mu_l + \mu_h)$. Hence, the match is uncertain only in the case that the worker changes firm after his starting firm from which he receives training.

Second, there are search frictions. I assume a constant returns to scale matching function, $m(v_i, n)$, where n is the number of workers, who can be either unemployed or employed, and v_i is the number of vacancies created by firm of type i, where $i \in \{a, b, c\}$. A worker, employed or unemployed, meets a vacant job in firm i with probability $p_i = m/n$. A vacancy meets a worker with probability $q_i = m/v_i$. Constant returns to scale implies that p_i and q_i are functions only of the ratio v_i/n. I also assume two boundary conditions: as $v_i/n \to \infty$, $p_i \to 1$ and $q_i \to 0$, and as $v_i/n \to 0$, $p_i \to 0$ and $q_i \to 1$.

The unit of production is a matched worker-firm pair. The observed output of a match is given by:

$$\tilde{y}_{ij} = y_{ij} + \varepsilon$$

for firm type $i \in \{a, b, c\}$ and for any worker who received training at a different starting firm, $j \neq i$. True match quality is y_{ij} and, depending on the hiring firm-source firm combination, match quality is drawn from a normal distribution with mean μ_τ and variance σ^2, $\tau \in \{l, h\}$; ε is a mean-zero independently and identically distributed random variable, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. The role of the "noise" term ε is to prevent firms and workers from perfectly inferring match quality immediately after first producing.

There are two different albeit related learning processes. The first concerns learning about the idiosyncratic match quality and evolves according to the firm-worker pair. If a worker and firm decide to form a match, they attempt to infer its quality from the observed match outputs. Learning takes a simple "all-or-nothing" form: each period, firms and workers fully learn about the worker's true productivity with probability α. The quality of a match is persistent – a good match remains good and a bad match remains bad – as long as it remains intact given the exogenous job destruction rate, δ. Moreover, firms do not know if the distribution of match qualities has a low or a high mean. This second type of learning, about the mean of the distribution, is not match-specific but it is source firm-specific: it is a belief on the average match quality of all the workers hired from the same starting firm. Firm i has a prior belief on the probability of the mean being high, $Prob(\mu_{ij} = \mu_h)$, which it updates to π_{ij} using the true match qualities, once they are fully revealed, and Bayes' rule.

Hence, there are two sources of uncertainty: the quality of the specific match and the value of the mean of its distribution, which depends on the hiring firm-source firm combination.

First, firms and workers use the signal, $\tilde{y}_{ij}$, to update their belief about the worker's idiosyncratic productivity. Denote the updated belief of the match quality as $m_{ij} = E(y_{ij}|\tilde{y}_{ij}, \pi_{ij})$, where π_{ij} is the employer's belief on the probability that the mean of the distribution is high. Denote the distribution of expected match qualities as $G(m_{ij}|\pi_{ij})$. Note that:[5]

$$m_{ij}|\pi_{ij} \sim \pi_{ij} N\left(\mu_h, \frac{\sigma^4}{\sigma^2 + \sigma_\varepsilon^2}\right) + (1 - \pi_{ij}) N\left(\mu_l, \frac{\sigma^4}{\sigma^2 + \sigma_\varepsilon^2}\right) \tag{3.4}$$

[5] See the derivations in Appendix A.

Hence, expected match qualities follow a mixture distribution of two normals with mean:

$$E(m_{ij}|\pi_{ij}) = \tilde{\mu} = \pi_{ij}\mu_h + (1-\pi_{ij})\mu_l \tag{3.5}$$

and variance:

$$V(m_{ij}|\pi_{ij}) = \frac{\sigma^4}{\sigma^2+\sigma_\varepsilon^2} + \pi_{ij}\mu_h^2 + (1-\pi_{ij})\mu_l^2 - (\pi_{ij}\mu_h + (1-\pi_{ij})\mu_l)^2 \tag{3.6}$$

Moreover, firms learn about the probability that the mean of the distribution of the match qualities with workers coming from a particular source firm is high. This belief on the mean, instead of being match-specific, is about the hiring firm-source firm combination. It is set on a prior of $Prob(\mu_{ij} = \mu_h) = 0.5$ and is updated to π_{ij} using Bayes' rule once true match qualities are fully revealed. Note that the mean is a discrete random variable: it can take only two values, $\mu \in \{\mu_l, \mu_h\}$. On the other hand, true match quality, y, is a continuous random variable taken from the normal distribution, $N(\mu_\tau, \sigma^2)$, $\tau \in \{l, h\}$. Since the belief about the mean takes into account the realizations of match quality for all workers hired from a specific source firm, Bayes' rule is computed using the mean of the realizations, $\bar{y}_{ij} \sim N(\mu_\tau, \sigma^2/n_{\bar{y}})$, $\tau \in \{l, h\}$ where $n_{\bar{y}}$ is the total number of matches whose true quality is revealed. Hence, as the true productivity of a worker hired from firm j is revealed, firm i updates its belief to:

$$\pi_{ij} = Prob(\mu_{ij} = \mu_h|\bar{y}_{ij}) = \frac{Prob(\mu_{ij} = \mu_h)f(\bar{y}_{ij}|\mu_{ij} = \mu_h)}{\sum_{\mu'} Prob(\mu_{ij} = \mu')f(\bar{y}_{ij}|\mu_{ij} = \mu')} \tag{3.7}$$

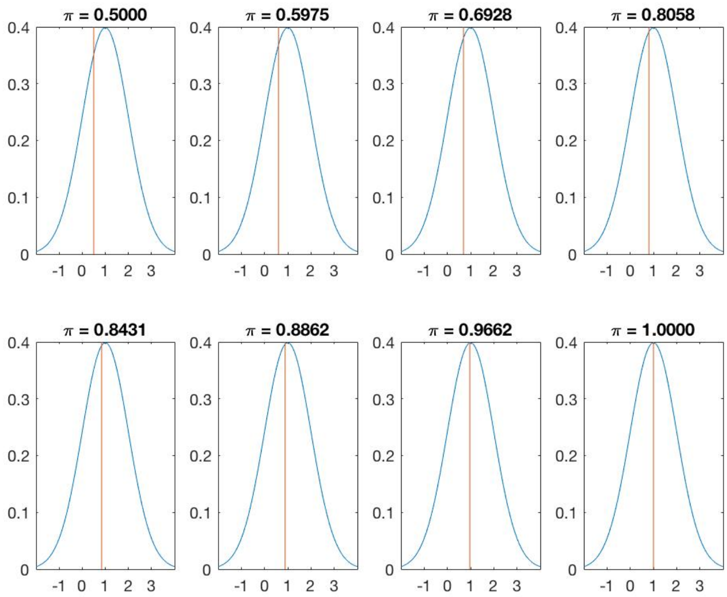
where $Prob(\mu_{ij} = \mu_h)$ is the unconditional probability that the mean is high and it is equal to 0.5 at the start and $f(\bar{y}_{ij}|\mu_{ij} = \mu_h)$ is the probability density function of the sample mean of the realizations of true match qualities given the mean is high, $f(\bar{y}_{ij}|\mu_{ij} = \mu_h) = N(\mu_h, \sigma^2/n_{\bar{y}})$. Figure 3.7 graphically shows how the belief about the mean is updated as the firm learns the true match quality of one additional worker each period.

Let $F(y|m_{ij}, \sigma^2_{ij,1}, \pi_{ij})$ denote the distribution of the worker's productivity y_{ij}, given that the signal is $\tilde{y}_{ij}$ and that the belief that the mean is high is π_{ij}. Since $F(\cdot)$ is a mixture of two normal distribution it can be characterized by the three parameters of the mean, m_{ij}, the variance, $\sigma^2_{ij,1}$ and the mixing probability, π_{ij}.

Using the formula for the conditional Gaussian model and considering that $F(\cdot)$ is a mixture distribution I obtain that[6] (Brockwell and Davis, 2016):

$$F(y|m_{ij}, \sigma^2_{ij,1}, \pi_{ij}) = \pi_{ij} N\left(\frac{\mu_h\sigma_\varepsilon^2 + \tilde{y}\sigma^2}{\sigma^2+\sigma_\varepsilon^2}, \frac{\sigma^2\sigma_\varepsilon^2}{\sigma^2+\sigma_\varepsilon^2}\right) + (1-\pi_{ij}) N\left(\frac{\mu_l\sigma_\varepsilon^2 + \tilde{y}\sigma^2}{\sigma^2+\sigma_\varepsilon^2}, \frac{\sigma^2\sigma_\varepsilon^2}{\sigma^2+\sigma_\varepsilon^2}\right) \tag{3.8}$$

[6] See the derivations in Appendix A.



Note: The plots show the pdf of a normal distribution with $\mu_h = 1$ and $\sigma^2 = 1$, $N(1, 1)$. The firm learns if the mean is high, $\mu_h = 1$, or low, $\mu_l = 0$, where π is the belief that the mean is high. The vertical line represents the expectation of the mean conditional on π, where π is updated using Equation 3.7 and the firm learns the true productivity of one new worker each period, i.e $n_{\bar{y}} = 1$. Above each plot, it is indicated the value of π that corresponds to the conditional expectation. The initial value of π is taken to be equal to 0.5.

Figure 3.7: Learning about the Mean

Hence the mean is:

$$m_{ij} = E(y_{ij}|\tilde{y}_{ij}, \pi_{ij}) = \pi_{ij}\frac{\mu_h\sigma^2 \;\; + \tilde{y}\sigma^2}{\sigma^2 + \sigma^2} + (1 - \pi_{ij})\frac{\mu_l\sigma^2 \;\; + \tilde{y}\sigma^2}{\sigma^2 + \sigma^2} \tag{3.9}$$

and the variance is:

$$\begin{aligned}\sigma^2_{ij,1} = Var(y_{ij}|\tilde{y}_{ij}, \pi_{ij}) = \frac{\sigma^2\sigma^2_\varepsilon}{\sigma^2 + \sigma^2_\varepsilon} + \pi_{ij}\left(\frac{\mu_h\sigma^2_\varepsilon + \tilde{y}\sigma^2}{\sigma^2 + \sigma^2_\varepsilon}\right)^2 + (1 - \pi_{ij})\left(\frac{\mu_l\sigma^2_\varepsilon + \tilde{y}\sigma^2}{\sigma^2 + \sigma^2_\varepsilon}\right)^2 \\ - \left(\pi_{ij}\left(\frac{\mu_h\sigma^2_\varepsilon + \tilde{y}\sigma^2}{\sigma^2 + \sigma^2_\varepsilon}\right) + (1 - \pi_{ij})\left(\frac{\mu_l\sigma^2_\varepsilon + \tilde{y}\sigma^2}{\sigma^2 + \sigma^2_\varepsilon}\right)\right)^2\end{aligned} \tag{3.10}$$

Wage Bargaining

In the case of employed workers, any new firm a worker comes in contact with perfectly observes the wage and the expected productivity of his current match and it receives a signal about the productivity of the potential match. The two firms play an English first-price auction and the auction ends after a firm fails to raise the last bid. The worker then receives the highest bid (lump-sum transfer) in exchange for giving up the contract with the losing firm and restarting Nash bargaining with the winner.

Hence, if a firm meets an employed worker, I use a generalized Nash bargaining solution for the wage in which the worker's threat point is the value of being unemployed, the firm's threat point is the value of an unmatched employment position, and the worker's share of the surplus is γ. Given the lump sum transfers by firms and the fact that on-the-job search is costless to the worker, wages set by the Nash bargaining linear sharing rule still satisfy the Nash efficiency axiom (Moscarini, 2005).

In the case of unemployed workers, I assume that firms have all the bargaining power and thus that the wage offered to an unemployed worker reflects just the value of leisure. The assumption that unemployed workers have no bargaining power may seem restrictive. However, there are at least two reasons that justify the choice. First, I assume that part of the worker's gain from being employed is to receive on-the-job training. Workers in the unemployment pool are inexperienced and a job can train them to then have the opportunity of better job prospects. Secondly, workers that are hired from the unemployment pool can potentially be poached away by competitors so that the employer suffers an unrecoverable loss. Giving all the bargaining power to the firm attenuates this cost.

Timing

The timing of the events in each period is as follows. Each firm can hire workers both from the unemployment pool and from the other firms. Workers start unemployed and meet jobs from any of the three firms. In this case, match quality is known, the firm offers a wage that makes workers indifferent between accepting the job and staying unemployed and workers accept the job offer.

In case the worker is poached from another firm, the poaching firm and the worker observe a signal about the quality of the match and the firm makes a wage offer. If the worker accepts, in the next period, the worker and the firm learn the productivity of the match with probability α and the firm updates its belief about the mean of the distribution of match qualities for the workers from a given source firm. The firm then makes a new wage offer and, if the employee turns it down, he becomes unemployed, and the position becomes vacant. I assume that, if a worker becomes unemployed, the skills learned at his training firm are lost and the worker would need to be trained again. Finally, each period, every match can be hit by a low productivity shock with probability δ, which will destroy it.

Starting Firm

A worker starts unemployed and can find a job at one of the three firms where he receives on the job training, match quality is $\left(0.5\mu_l + 0.5\mu_h\right)$ and there is no match uncertainty. The worker at starting firm i earns a wage w_{ii} and, with probability p_j, he receives an alternative job offer from firm $j \neq i$. The value function of a worker in the starting firm is:

$$\begin{aligned} W_{ii} = w_{ii} + & \beta(1-\delta)\Big(p_j E\, max\left(W_{ji,1}, W_{ii}\right) + p_k E\, max\left(W_{ki,1}, W_{ii}\right) + (1-p_j-p_k)W_{ii}\Big) \\ & + \beta\delta U \end{aligned} \tag{3.11}$$

for $i \in \{a, b, c\}$ and $i \neq j \neq k$. The notation is such that W_{ii} stands for the value function of a worker trained at firm i and currently employed at firm i and $W_{ji,1}$ stands for the value function of a worker being poached by firm j from firm i at the beginning of the employment spell.

The firm with an inexperienced worker earns $\left(0.5\mu_l + 0.5\mu_h\right) - w_{ii}$ and can lose the worker if he is poached by another firm. The value function of a firm with a worker hired from the unemployment pool is:

$$J_{ii} = \left(0.5\mu_l + 0.5\mu_h\right) - w_{ii} + \beta(1-\delta)\left(1 - p_j 1\{W_{ji,1} > W_{ii}\} - p_k 1\{W_{ki,1} > W_{ii}\}\right) J_{ii} \tag{3.12}$$

where $i \in \{a, b, c\}$ and 1 is an indicator function so $1\{W_{ji,1} > W_{ii}\} = 1$ only if a worker accepts an outside offer from firm j. I used the fact that free entry drives the value of vacancy to zero, $V_i = 0$.

The value function for being unemployed is:

$$U = b + \beta p_a \, max\,(W_{aa}, U) + \beta p_b \, max\,(W_{bb}, U) + \beta p_c \, max\,(W_{cc}, U) \qquad (3.13)$$
$$+ \beta(1 - p_a - p_b - p_c)U$$

where b is the unemployment benefit.

Firms post vacancies which can be filled either by employed workers or by unemployed ones. The value of a vacancy is given by:

$$V_i = -k_v + \beta \left(q_i \frac{n_j}{n} E \max(J_{ij,1}, V_i) + q_i \frac{n_k}{n} E \max(J_{ik,1}, V_i) + q_i \frac{u}{n} E \max(J_{ii}, V_i) \right) \qquad (3.14)$$
$$+ \beta(1 - q_i \frac{n_j + n_k + u}{n}) V_i$$

where $i \in \{a, b, c\}$, $i \neq j \neq k$ and n_i is the number of workers employed at firm i.

I assume that when firms meet an unemployed worker, they have all the bargaining power and thus that the wage offered to an unemployed worker reflects just the value of leisure. The wage will be that which equals the value of working and the value of being unemployed, $W_{ii} = U$.

Poaching firm

I now consider the value functions in the case a worker is poached from his starting firm by another firm. In this case, match quality is uncertain and it is drawn from a normal distribution, which, depending on the hiring firm-source firm pair, has a high or low mean, $\mu \in \{\mu_l, \mu_h\}$. Given the initial signal about match quality, a worker can choose to leave his starting firm to work at the new firm. Once the match quality is revealed, if too low, the worker chooses to quit the job and join the unemployment pool and, if high enough, the worker continues working at the firm. At this point, working at the poaching firm after the match quality is revealed is an absorbing state: the worker works at the firm as long as the job is not destroyed and there is no possibility of changing firm again without first becoming unemployed. Hence, a worker can be poached away by another firm only if he is working at the firm that hired him from the unemployment pool and trained him.

Poaching auction

When the employed worker comes into contact with a new firm, the two firms immediately play the English first-price auction and the auction ends after a firm fails to raise the last bid. Let Ω_{ij} denote the final bid of poaching firm i against firm j. Necessarily, $\Omega_{ij} \in [W_{ij,1}, S_{ij,1}]$ for the bid to be acceptable both to the bidding firm and to the worker, where $W_{ij,1}$ is the value of working at firm i having previously worked at firm j, given that match quality has not been revealed yet, and $S_{ij,1}$ denotes the surplus of the match, given that match quality has not been revealed yet. The bid is the sum of a lump-sum transfer $\omega_{ij} = \Omega_{ij} - W_{ij,1}$ and of the promise, worth $W_{ij,1}$ to the worker, to match, produce output and bargain bilaterally from then on. The auction is a symmetric information game in extensive form and the following strategies are a subgame perfect equilibrium of this game: each firm bids just the bargaining value $\Omega_{ij} = W_{ij,1}$ and no lump-sum transfer $\omega_{ij} = 0$. This equilibrium is supported by the threat of the more productive of the two firms to outbid the competitor at the next round (Moscarini, 2005).

Stage 2

I start with the value functions for being in a match with the poaching firm, once the true match quality has been realized. With probability $(1-\delta)$, the match survives and the value of the match remains unchanged. With probability δ, the job is destroyed for exogenous reasons. The worker's and firm's value functions, $W_{ij,2}$ and $J_{ij,2}$ are:

$$W_{ij,2} = w_{ij,2} + \beta(1-\delta)W_{ij,2} + \beta\delta U \tag{3.15}$$

$$J_{ij,2} = y_{ij} - w_{ij,2} + \beta(1-\delta)J_{ij,2} \tag{3.16}$$

where $w_{ij,2}$ is the wage offered to the worker, $W_{ij,2}$ indicates the value function of a worker poached by firm i from firm j when the productivity of the match has been fully revealed and $J_{ij,2}$ is defined similarly with respect to the firm's side.

Wages are determined by:

$$W_{ij,2} - U = \gamma(W_{ij,2} - U + J_{ij,2}) \tag{3.17}$$

where I use the fact that $V_i = 0$ due to free entry. There is a reservation match quality, y^* such that, if $y > y^*$, the worker prefers to stay at the firm and the firm chooses to keep the worker.

Stage 1

Each firm can meet workers that come from two distinct source firms. For example, firm a can meet a worker from firm b or from firm c. Considering firm a, the mean of the distribution of workers from firm b is μ_h while it is μ_l for workers from firm c.

As they meet, the worker from firm j and firm i observe a signal and compute the expected productivity of the match, m_{ij}. If hired, the worker earns $w_{ij,1}$ in the current period. Next period, the job is destroyed for exogenous reasons with probability δ. With probability $(1-\delta)(1-\alpha)$, the job survives, the firm and the worker receive no new information about the worker's productivity, and the worker's value of the match remains unchanged. Finally, with probability $\alpha(1-\delta)$, the job survives and the worker's productivity is revealed. In this case, the worker chooses between staying at the current firm or quitting the job and becoming unemployed. The worker's value of the match is given by:

$$\begin{aligned} W_{ij,1} = w_{ij,1} + &\beta(1-\alpha)(1-\delta)W_{ij,1} \\ &+ \beta\alpha(1-\delta)\int max(W_{ij,2}, U)dF(y|m_{ij}, \sigma^2_{ij,1}, \pi_{ij}) + \beta\delta U \end{aligned} \tag{3.18}$$

where $i \in \{a, b, c\}$ is the firm where the worker is currently employed and $j \neq i$ is the firm where he was trained.

The firm's value of the match is:

$$\begin{aligned} J_{ij,1} = m_{ij} - w_{ij,1} + &\beta(1-\alpha)(1-\delta)J_{ij,1} \\ &+ \beta\alpha(1-\delta)\int max(J_{ij,2}, 0)dF(y|m_{ij}, \sigma^2_{ij,1}, \pi_{ij}) \end{aligned} \tag{3.19}$$

Wages are determined by Nash bargaining where the worker's outside option is the value of unemployment:

$$W_{ij,1} - U = \gamma(W_{ij,1} - U + J_{ij,1}) \tag{3.20}$$

There is a reservation expected match quality, m^*_{ij} such that, if $m_{ij} > m^*_{ij}$, the worker prefers to stay at the new firm and the firm chooses to hire the worker.

3.3.2 Equilibrium

Reservation match qualities

I solve for the reservation match qualities and reservation expected match qualities for the poached workers in Stage 2 and Stage 1.

Considering first the reservation match qualities (Stage 2), workers stay with the firm if the total surplus of the match $S_{ij,2} = W_{ij,2} - U + J_{ij,2}$ is positive. Rearranging $W_{ij,2}$ and $J_{ij,2}$ and adding them up yields:

$$S_{ij,2} = \frac{y_{ij} - (1-\beta)U}{1-\beta(1-\delta)}$$

It follows that the reservation match quality is:

$$y^* = (1-\beta)U \tag{3.21}$$

Note that the reservation match quality does not depend on the hiring firm-source firm pair but only depends on the value of being unemployed.

Next, I consider the reservation expected match qualities (Stage 1). Workers choose to move to the poaching firm if the total surplus of the match $S_{ij,1} = W_{ij,1} - U + J_{ij,1}$ is positive[7]. Rearranging $W_{ij,1}$ and $J_{ij,1}$, adding them up and using $y^* = (1-\beta)U$, yields:

$$S_{ij,1} = \frac{m_{ij} - y^* + \frac{\alpha\beta(1-\delta)}{1-\beta(1-\delta)} \int_{y^*}^{\infty} (y_{ij} - y^*) dF(y|m_{ij}, \sigma^2_{ij,1}, \pi_{ij})}{1 - \beta(1-\alpha)(1-\delta)}$$

Hence, the reservation expected match quality is:

$$m^*_{ij} = y^* - \frac{\alpha\beta(1-\delta)}{1-\beta(1-\delta)} \int_{y^*}^{\infty} (y_{ij} - y^*) dF(y|m_{ij}, \sigma^2_{ij,1}, \pi_{ij}) \tag{3.22}$$

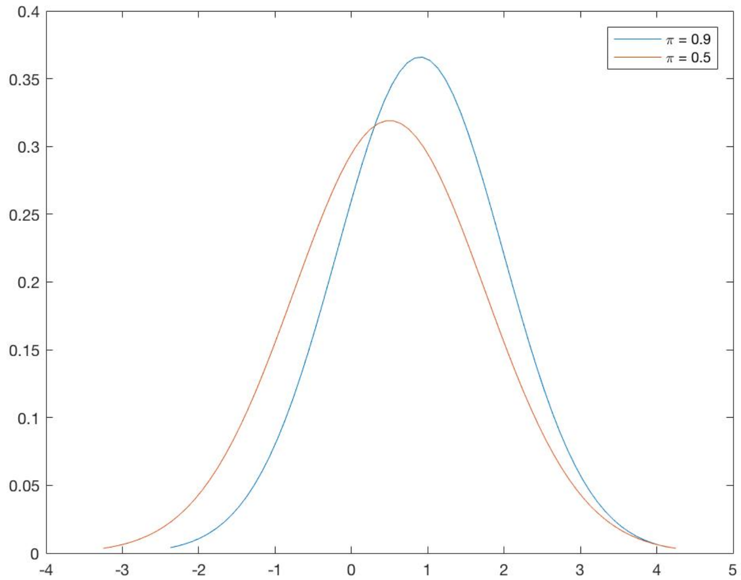
Note that the reservation expected match quality depends on both the hiring firm and the source firm.

Learning about the mean

To better understand how learning about the mean of the distribution interacts with learning about the idiosyncratic match quality, I establish some results, which follow from the formulas of the distributions. Figure 3.8 shows the mixture of two normal distributions for different values of the mixing probability and graphically illustrates the points below.

- The higher the belief that the mean is high, π_{ij}, the higher the conditional expectation of match qualities, $E(y_{ij}|\tilde{y}_{ij}, \pi_{ij}) = m_{ij}$.
- The more uncertain the belief that the mean is high, i.e. the closer π_{ij} is to 0.5, the higher the conditional variance of match qualities, $V(y_{ij}|\tilde{y}_{ij}, \pi_{ij})$.
- The higher the belief that the mean is high, π_{ij}, the higher the expectation of expected match quality, $E(m_{ij}|\pi_{ij})$.
- The more uncertain the belief that the mean is high, i.e. the closer π_{ij} is to 0.5, the higher the variance of the expected match qualities, $V(m_{ij}|\pi_{ij})$.
- The higher the belief that the mean of the distribution is high, π_{ij}, the lower the reservation expected match quality, m^*_{ij}. The intuition follows from the fact that firms and workers are willing to accept worse matches the greater the expectation that the mean of the distribution is high. In fact, with the additional learning process about the mean, the firm relies less on the signal about the idiosyncratic match quality, $\tilde{y}_{ij}$. For instance,

[7] I have assumed that the wage in the starting firm is such that $W_{ii} = U$.



Note: The figure shows the probability density function of the mixture of two normal distributions, $\pi N(1, 1) + (1 - \pi)N(0, 1)$, for different values of the mixing probability, $\pi = 0.9$ or $\pi = 0.5$. When $\pi = 0.5$, the distribution is shifted to the left (lower mean) and has fatter tails (higher variance).

Figure 3.8: Mixture of Two Normal Distributions, Different Values of π

assuming the firm knows that there is a high probability that the mean is high, a low signal may just be due to the noise term. Similarly, the lower the belief that the mean of the distribution is high, π_{ij}, the higher the reservation expected match quality. In this case, the firm requires a high signal for the quality of the match since there is a plausible chance that true match quality will be low.

- The reservation match quality, y^*, does not depend on the belief that the mean is high, π_{ij}, since at that point the productivity of the match is fully revealed.

Steady-state flows

In steady state, the outflow of workers from firm j is equal to the inflow of workers into firm j, $j \in \{a, b, c\}$, so that each firm has a constant fraction of the total workers. Note that, in steady state, each firm has fully learned about the mean of the distributions of the workers hired from its source firms and only the first learning process, about the idiosyncratic match quality, remains active.

For any firm i, the outflow of workers is equal to:

$$outflow_i = \delta n_i + \underbrace{(1-\delta)p_j(1-G(m^*_{ji}|\pi_{ji}))n_i}_{n_{ji}} + \underbrace{(1-\delta)p_k(1-G(m^*_{ki}|\pi_{ki}))n_i}_{n_{ki}} \quad (3.23)$$
$$+ \alpha(1-\delta)\sum_{j\neq i}\left(\int_{m^*_{ij}}^{\infty}\int_{-\infty}^{y^*}\frac{dF(y|m_{ij},\sigma^2_{ij,1},\pi_{ij})dG(m_{ij}|\pi_{ij})n_{ij}}{1-G(m^*_{ij}|\pi_{ij})}\right)$$

where $j \neq k \neq i$, and n_i is the total number of workers employed at firm i.

The inflow of workers is equal to:

$$inflow_i = p_i u + \underbrace{(1-\delta)p_i(1-G(m^*_{ij}|\pi_{ij}))n_j}_{n_{ij}} + \underbrace{(1-\delta)p_i(1-G(m^*_{ik}|\pi_{ik}))n_k}_{n_{ik}} \quad (3.24)$$

The outflows from unemployment are:

$$outflows_u = pu \quad (3.25)$$

where $p = p_a + p_b + p_c$.

The inflows into unemployment are given by:

$$inflows_u = \delta(n_a + n_b + n_c) \quad (3.26)$$
$$+ \alpha(1-\delta)\sum_{i}\sum_{j\neq i}\left(\int_{m^*_{ij}}^{\infty}\int_{-\infty}^{y^*}\frac{dF(y|m_{ij},\sigma^2_{ij,1},\pi_{ij})dG(m_{ij}|\pi_{ij})n_{ij}}{1-G(m^*_{ij}|\pi_{ij})}\right)$$

By equating outflows (Equation 3.23) and inflows (Equation 3.24) respectively for each firm I obtain the number of workers in each firm, $n_i,\ i \in \{a, b, c\}$. The number of unemployed workers is given by equating Equation 3.25 with Equation 3.26. Finally, the number of vacancies for each firm, $v_i,\ i \in \{a, b, c\}$, is pinned down by the free entry condition and Equation 3.14.

Wages

Using Equations 3.11-3.13 and $W_{ii} = U$, the average wage of a worker hired from the unemployment pool is:

$$\begin{aligned}\bar{w}_{ii} &= (1-\beta(1-\delta)(1-p_j-p_k)-\beta\delta)\frac{b}{1-\beta} \\ &- \beta(1-\delta)\gamma\Bigg(p_j\Big(\frac{\int_{m_{ji}^*}^{\infty}\big(m_{ji}-y^*+\frac{\alpha\beta(1-\delta)}{1-\beta(1-\delta)}\int_{y^*}^{\infty}(y_{ji}-y^*)dF(y|m_{ji},\sigma_{ji,1}^2,\pi_{ji})\big)dG(m_{ji}|\pi_{ji})}{1-\beta(1-\alpha)(1-\delta)}\Big) \\ &+ p_k\Big(\frac{\int_{m_{ki}^*}^{\infty}\big(m_{ki}-y^*+\frac{\alpha\beta(1-\delta)}{1-\beta(1-\delta)}\int_{y^*}^{\infty}(y_{ki}-y^*)dF(y|m_{ki},\sigma_{ki,1}^2,\pi_{ki})\big)dG(m_{ki}|\pi_{ki})}{1-\beta(1-\alpha)(1-\delta)}\Big)\Bigg)\end{aligned} \tag{3.27}$$

Using Equations 3.18-3.20, the average wage of the workers in the poaching firm whose productivity has not been revealed yet is given by:

$$\begin{aligned}\bar{w}_{ij,1} &= \frac{\int_{m_{ij}^*}^{\infty} w_{ij,1}dG(m_{ij}|\pi_{ij})}{1-G(m_{ij}^*|\pi_{ij})} \\ &= \frac{\gamma\int_{m_{ij}^*}^{\infty} m_{ij}dG(m_{ij}|\pi_{ij})}{1-G(m_{ij}^*|\pi_{ij})} + (1-\beta)(1-\gamma)U\end{aligned} \tag{3.28}$$

Using Equations 3.15-3.17, the average wage of the workers whose productivity has been revealed:

$$\begin{aligned}\bar{w}_{ij,2} &= \frac{\int_{m_{ij}^*}^{\infty}\frac{\int_{y^*}^{\infty} w_{ij,2}dF(y|m_{ij},\sigma_{ij,1}^2,\pi_{ij})dG(m_{ij}|\pi_{ij})}{1-F(y^*|m_{ij},\sigma_{ij,1}^2,\pi_{ij})}}{1-G(m_{ij}^*|\pi_{ij})} \\ &= \frac{\gamma\int_{m_{ij}^*}^{\infty}\frac{\int_{y^*}^{\infty} y_{ij}dF(y|m_{ij},\sigma_{ij,1}^2,\pi_{ij})dG(m_{ij}|\pi_{ij})}{1-F(y^*|m_{ij},\sigma_{ij,1}^2,\pi_{ij}),\pi_{ij}}}{1-G(m_{ij}^*|\pi_{ij})} + (1-\beta)(1-\gamma)U\end{aligned} \tag{3.29}$$

Turnover

The worker can leave his current firm for three different reasons. First, the worker leaves the starting firm in case of an alternative offer from a poaching firm. Second, the worker leaves the poaching firm if, upon the discovery of the true match quality, the productivity of the match is too low. Finally, in both cases, the job can be destroyed, in which case the worker becomes unemployed.

The probability that a worker leaves his starting firm next period is:

$$Prob_{ii} = \delta + (1-\delta)\left(p_j\int_{m_{ji}^*}^{\infty} dG(m_{ji}|\pi_{ji}) + p_k\int_{m_{ki}^*}^{\infty} dG(m_{ki}|\pi_{ki})\right) \tag{3.30}$$

where $i \neq j \neq k$.

The probability that a worker whose productivity has not been revealed yet leaves the poaching firm in the next period is:

$$Prob_{ij} = \delta + \frac{\alpha(1-\delta)\int_{m_{ij}^*}^{\infty}\int_{-\infty}^{y^*} dF(y|m_{ij}, \sigma^2_{ij,1}, \pi_{ij})dG(m_{ij}|\pi_{ij})}{\int_{m_{ij}^*}^{\infty} dG(m_{ij}|\pi_{ij})} \tag{3.31}$$

The probability that a worker whose productivity has been revealed leaves the poaching firm in the next period is δ.

3.4 Quantitative Analysis

3.4.1 Calibration

I choose the time period to be one year. I assume a constant returns to scale matching function, $m(n, v_i) = n^{\eta} v_i^{1-\eta}$. I normalize the mean of the low productivity distribution to $\mu_l = 0$. I select the value of μ_h to minimize the squared distance between the model and the data moment of the starting wage differential depending on the hiring experience of the firm (see Figure 3.3 in the empirical section).[8] The variance of productivity, σ^2, and of the noise, σ^2_{ε}, are taken from Nagypál (2007), who estimates the parameters based on a structural model using French data. The learning rate, α, is taken from Lange (2007), who empirically estimates the speed of employer learning using US data. The remaining exogenous parameters are taken from Borovicková (2016), who estimates an equilibrium model of the labor market with Austrian data.[9] I follow the literature and set workers' bargaining power, γ, equal to the match elasticity (Pries and Rogerson, 2005). Table 3.8 lists the values of the exogenous parameters of the model.

To compute the model moments, I first compute the endogenous variables, y^*, m_{ij}^*, n_i, u, v_i, $i \in \{a, b, c\}$, of the model for the parameter values in Table 3.8.

These are the Equations that determine these variables:

- The reservation expected match qualities of workers, m_{ij}^*, are given by Equation 3.22.
- The reservation match quality of workers, y^*, is given by Equation 3.21. Plugging in the value of unemployment and noting that wages are set such that, $W_{ii} = U$:

$$\begin{aligned} y^* &= (1-\beta)U \\ &= b \end{aligned}$$

[8] I consider the difference in the log daily wages between the case of a hiring experience of 100 previous workers compared to the case of no previous hirings (Figure 3.3).

[9] Borovicková (2016) uses the same data of this paper, the Austrian Social Security Database (ASSD).

Table 3.8: Calibration

Parameter	Description	Value	Source
β	Discount factor	0.9542	Borovicková (2016)
b	Unemployment benefit	0.24	Borovicková (2016)
k_v	Cost vacancy		Calibrated to match steady-state unemployment, 0.12 (Borovicková, 2016)
η	Match elasticity	0.5	Borovicková (2016)
α	Rate at which match quality is revealed	0.2592	Lange (2007)
δ	Rate of exogenous job destruction	0.1032	Borovicková (2016)
γ	Worker's bargaining power	0.5	Hosios condition (Hosios, 1990)
μ_l	Low Mean Productivity	0	Normalization
σ^2	Variance of Productivity	0.3920	Nagypál (2007)
σ^2	Variance of the noise	1.0574	Nagypál (2007)

- The free entry condition and Equation 3.14:

$$k_v = \beta(1-\gamma)q_i\left(\frac{n_j}{s}\left(\frac{\int_{m_{ij}^*}^{\infty}\frac{m_{ij}-y^*+\frac{\alpha\beta(1-\delta)}{1-\beta(1-\delta)}\int_{y^*}^{\infty}(y_{ij}-y^*)dF(y|m_{ij},\sigma_{ij,1}^2,\pi_{ij})dG(m_{ij}|\pi_{ij})}{1-G(m_{ij}^*|\pi_{ij})}}{1-\beta(1-\alpha)(1-\delta)}\right)\right.$$
$$+\frac{n_k}{s}\left(\frac{\int_{m_{ik}^*}^{\infty}\frac{m_{ik}-y^*+\frac{\alpha\beta(1-\delta)}{1-\beta(1-\delta)}\int_{y^*}^{\infty}(y_{ik}-y^*)dF(y|m_{ik},\sigma_{ik,1}^2,\pi_{ik})dG(m_{ik})}{1-G(m_{ik}^*|\pi_{ik})}}{1-\beta(1-\alpha)(1-\delta)}\right)$$
$$\left.+\frac{u}{s}\left(\frac{0.5\mu_l+0.5\mu_h-w_{ii}}{1-\beta(1-\delta)(1-p_j(1-G(m_{ji}^*|\pi_{ji}))-p_k(1-G(m_{ki}^*|\pi_{ki})))}\right)\right)$$

  where the wage, w_{ii}, is given by Equation 3.27.

- The equality of the outflow of workers out of unemployment, given by Equation 3.25, and the inflow of workers into unemployment, given by Equation 3.26:

$$(p_a+p_b+p_c)u = \delta(n_a+n_b+n_c)$$
$$+\alpha(1-\delta)\sum_i\sum_{j\neq i}\left(\int_{m_{ij}^*}^{\infty}\int_{-\infty}^{y^*}\frac{dF(y|m_{ij},\sigma_{ij,1}^2,\pi_{ij})dG(m_{ij}|\pi_{ij})n_{ij}}{1-G(m_{ij}^*|\pi_{ij})}\right)$$

- The equality of the outflow of workers out of a firm, given by Equation 3.23, and the inflow of workers into a firm, given by Equation 3.24:

$$\begin{aligned}
&\delta n_i + \underbrace{(1-\delta)p_j(1-G(m^*_{ji}|\pi_{ji}))n_i}_{n_{ji}} + \underbrace{(1-\delta)p_k(1-G(m^*_{ki}|\pi_{ki}))n_i}_{n_{ki}} \\
&+ \alpha(1-\delta)\sum_{j\neq i}\left(\int_{m^*_{ij}}^{\infty}\int_{-\infty}^{y^*}\frac{dF(y|m_{ij},\sigma^2_{ij,1},\pi_{ij})dG(m_{ij}|\pi_{ij})n_{ij}}{1-G(m^*_{ij}|\pi_{ij})}\right) \\
&= p_i u + \underbrace{(1-\delta)p_i(1-G(m^*_{ij}|\pi_{ij}))n_j}_{n_{ij}} + \underbrace{(1-\delta)p_i(1-G(m^*_{ik}|\pi_{ik}))n_k}_{n_{ik}}
\end{aligned}$$

- The matching technology gives $p_j = (\frac{v_j}{n})^{(1-\eta)}$ and $q_j = (\frac{n}{v_j})^{\eta}$.

3.4.2 Quantitative Results

I am interested in analyzing the worker flows and the wage and turnover outcomes. Regarding the worker flows, I look at the proportion of the workers employed in a firm that is poached away from another firm. The number of workers employed in firm i poached from firm j, n_{ij}, is given by: $n_{ij} = (1-\delta)p_i(1-G(m^*_{ij}|\pi_{ij}))n_j$. Table 3.9 compares the results between the steady-state where the firm has fully learned about the mean of the distributions and the steady state where the learning mechanism is shut down and the firm believes there is an equal chance that the mean is high or low ($\pi = 0.5$). First, when $\pi = 0.5$, firms do not discriminate between their sources and they poach the same proportion of workers from all of them. However, when they fully learn, hiring gets more concentrated and, considering the network structure of Table 3.7, the results are intuitive. Firm a poaches more workers from firm b, with whom the distribution of match qualities has a high mean, compared to firm c, with whom the distribution of match qualities has a low mean. The ratio of the workers hired by firm a from firm b to the workers hired from firm c is above 2:1. Given the symmetric network structure, firm c is the analogue of firm a. Instead, firm b poaches the same amount of workers from firm a and firm c since both firms train their workers to acquire the skills that are highly valued by firm b.[10]

Secondly, I consider worker turnover. The probability that a worker leaves his starting firm next period is given by Equation 3.30. The probability of separation for a worker whose productivity has not been revealed yet is given by Equation 3.31. Once productivity is revealed, the worker leaves the firm only in case of job destruction, which occurs with probability δ and is the same for all workers regardless of their starting firm. In Table 3.9 I compute the probability of leaving the starting firm, $Prob_{ii},\ i \in \{a,b,c\}$, as well as the probability of leaving the poaching firm, given the productivity of the match has not been revealed yet, $Prob_{ij}, i \neq j$. By the symmetry of the network structure (c.f. Table 3.7), I have that $Prob_{ij} = Prob_{ji}$. In

[10] The distribution of match qualities has a high mean in both cases.

Table 3.9: Estimated Moments (Steady-state)

Moment	Value ($\pi = 0.5$)	Value
n_{ab}	33.27%	33.03%
n_{ac}	33.27%	11.35%
n_{ba}	33.27%	20.96%
n_{bc}	33.27%	20.96%
n_{ca}	33.27%	11.35%
n_{cb}	33.27%	33.03%
$Prob_{aa}$	0.7686	0.5401
$Prob_{bb}$	0.7686	0.5316
$Prob_{cc}$	0.7686	0.5401
$Prob_{ab} = Prob_{ba}$	0.1768	0.1047
$Prob_{ac} = Prob_{ca}$	0.1768	0.2678
$Prob_{bc} = Prob_{cb}$	0.1768	0.1047

fact, what discriminates between the different cases is whether the hiring firm-source firm combination is high or low.

When the firm does not know about the true mean of the distributions, the probability of leaving a starting or a poaching firm does not change across the different firms (Table 3.9). However, this is no longer the case when the firm knows the mean of the distributions. Considering the probability of leaving a poaching firm, it is more than two times more likely that a worker who is poached by firm a/c from firm c/a leaves the poaching firm compared to the other two combinations. Indeed, this is the case where the hiring firm-source firm combination has a low mean distribution of match qualities. Moreover, $Prob_{ab} = Prob_{ba} = Prob_{bc} = Prob_{cb}$ ($Prob_{ac} = Prob_{ca}$) is greater (lower) when $\pi = 0.5$ than when $\pi = 1$.

Turning to the probability of leaving the starting firm, I need to take into account two contrasting effects. First, the training received at firm b is valued both at firm a and c, which points to a higher chance that the workers trained at firm b leave the firm for a better job opportunity at either firm a or firm c. However, at the same time, firm b can have profitable matches with the workers trained both at firm a and at firm c, encouraging firm b to post a higher number of vacancies. In this case, the second effect dominates leading to a slightly higher probability of separation from firm a and c compared to firm b since workers from firm a and c have a greater chance of finding an alternative job opportunity given the relatively high vacancy posting of firm b. By the symmetry of the network structure, the probability of leaving firm a is the same as the probability of leaving firm c.

Finally, in all the hiring firm-source firm combinations, the probability that a worker leaves the poaching firm given that true productivity has not been revealed yet is higher than the probability that a worker leaves the poaching firm given that true productivity has been revealed, which, in all cases, is given by $\delta = 0.1032$.

Considering the wages, for each firm i hiring workers from firm $j \neq i$, I calculate the difference between the average wage when the firm has fully learned about the mean of the distribution of match qualities and the average wage when the firm does not know the true value of the mean. I compute the moment both at the beginning of the employment relationship, when true idiosyncratic productivity is unknown, $\Delta\bar{w}_{ij,1}$, and for the subsequent years of employment, $\Delta\bar{w}_{ij,av}$. In the first case, $\bar{w}_{ij,1}$ is given by Equation 3.28. The mean wage in subsequent years of the employment relationship, $\bar{w}_{ij,av}$, is a weighted average of the wage of workers whose productivity has not been revealed yet, $\bar{w}_{ij,1}$ (Equation 3.28), and that of workers whose productivity is known, $\bar{w}_{ij,2}$ (Equation 3.29). The moment is given by:

$$\bar{w}_{ij,av} = \frac{(1-\alpha)\bar{w}_{ij,1} + \frac{\alpha}{\delta}\frac{\int_{m^*_{ij}}^{\infty}(\int_{y^*}^{\infty} dF_{ij})dG_{ij}\bar{w}_{ij,2}}{\int_{m^*_{ij}}^{\infty} dG_{ij}}}{(1-\alpha) + \frac{\alpha}{\delta}\frac{\int_{m^*_{ij}}^{\infty}(\int_{y^*}^{\infty} dF_{ij})dG_{ij}}{\int_{m^*_{ij}}^{\infty} dG_{ij}}}$$

Moreover, I look at how the wage difference changes between the two periods. I expect $\Delta\bar{w}_{ij,1} - \Delta\bar{w}_{ij,av} > 0$, since, once the firm has learned about the idiosyncratic match quality, it matters less whether it knows about the mean of the distribution.

The different hiring firm-source firm combinations can be reduced to two cases: the ones with a high mean distribution of match qualities and those with a low mean distribution. In Table 3.10 I illustrate the average wage difference between the case when the firm has learned the value of the mean of the distributions of match qualities and the case when it has not ($\pi = 0.5$).

Table 3.10: Estimated Wage Differential

Moment	Value
$\Delta\bar{w}_{h,1}$	0.2486
$\Delta\bar{w}_{l,1}$	-0.2116
$\Delta\bar{w}_{h,av}$	0.1718
$\Delta\bar{w}_{l,av}$	-0.2057
$\Delta\bar{w}_{h,1} - \Delta\bar{w}_{h,av}$	0.0768
$\lvert\Delta\bar{w}_{l,1}\rvert - \lvert\Delta\bar{w}_{l,av}\rvert$	0.0059

In the high (low) mean case, I find that average starting wages are higher (lower) when the firm knows that the mean is high (low).[11] Moreover, the wage difference declines over time as

[11] This result always holds although the reservation expected match qualities evolve in the opposite direction. For example, in the high mean case, the more the firm has learned about the value of the mean, the lower is the

the firm learns about the true productivity of the specific matches.

3.5 Conclusion

I explore the role of hiring experience a firm gains by repeatedly poaching workers from the same source firms. Using large-scale matched employer–employee data covering private sector employment in Austria, I show that more experienced firms hire their employees from a narrower set of source firms. Experienced hires lead to more stable employment relations and higher starting wages. Over time, the information advantage dissipates as firms and workers learn about actual match quality.

The empirical findings are rationalized by a model of job-to-job flows where match quality is uncertain and firms learn over time from which firms to source their employees. In the model, there are two different learning processes. The first is the well-known learning about the idiosyncratic match quality of the firm-worker pair. The second is the employer's learning about the mean of the distribution of match qualities for a particular hiring firm-source firm combination. The model framework assumes that firms are related to each other through a network of labor flows that expresses skill transferability. The model suggests a novel perspective on the role of employer's learning about its source firms in shaping the direction of labor flows and in improving the matching between workers and firms. The more a firm learns about its source firms, the better the matches of the new hires resulting in higher wages and lower turnover. However, the wage and turnover advantage dissipates over time as firms learn about match-specific productivity and bad matches are terminated.

reservation expected match quality. Nonetheless, starting wages are higher. The fact that the higher π shifts the distribution of expected match qualities to the right dominates unambiguously the fact that the higher π lowers the reservation expected match quality.

References

Joseph G Altonji and Robert A Shakotko. Do Wages Rise with Job Seniority? *The Review of Economic Studies*, 54(3):437–459, 1987.

Fredrik Andersson, Matthew Freedman, John Haltiwanger, Julia Lane, and Kathryn Shaw. Reaching for the Stars: Who Pays for Talent in Innovative Industries? *The Economic Journal*, 119(538):F308–F332, 2009.

Jesper Bagger, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics. *The American Economic Review*, 104(6):1551–1596, 2014.

Patrick Bayer, Stephen L Ross, and Giorgio Topa. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116 (6):1150–1196, 2008.

Thomas Behrends. Recruitment Practices in Small and Medium Size Enterprises. An Empirical Study among Knowledge-intensive Professional Service Firms. *Management Revue*, pages 55–74, 2007.

Annette Bergmann and Antje Mertens. Job Stability Trends, Lay-offs, and Transitions to Unemployment in West Germany. *Labour*, 25(4):421–446, 2011.

Katarına Borovicková. Job Flows, Worker Flows and Labor Market Policies. *Mimeo*, 2016.

Udo Brixy, Susanne Kohaut, and Claus Schnabel. Do Newly Founded Firms Pay Lower Wages? First Evidence from Germany. *Small Business Economics*, 29(1):161–171, 2007.

Peter J Brockwell and Richard A Davis. *Introduction to Time Series and Forecasting*. Springer, 2016.

Charles Brown and James Medoff. The Employer Size-wage Effect. *Journal of Political Economy*, 97(5):1027–1059, 1989.

Charles Brown and James L Medoff. Firm Age and Wages. *Journal of Labor Economics*, 21 (3):677–697, 2003.

Kenneth Burdett and Dale T. Mortensen. Wage Differentials, Employer Size, and Unemployment. *International Economic Review*, 39(2):257–273, 1998.

Simon Burgess, Julia Lane, and David Stevens. Hiring Risky Workers: Some Evidence. *Journal of Economics & Management Strategy*, 7(4):669–676, 1998.

Federico Cingano and Alfonso Rosolia. People I know: Job Search and Social Networks. *Journal of Labor Economics*, 30(2):291–332, 2012.

Jed DeVaro. Employer Recruitment Strategies and the Labor Market Outcomes of New Hires. *Economic Inquiry*, 43(2):263–282, 2005.

Jed DeVaro. The Labor Market Effects of Employer Recruitment Choice. *European Economic Review*, 52(2):283–314, 2008.

Christian Dustmann and Costas Meghir. Wages, Experience and Seniority. *The Review of Economic Studies*, 72(1):77–108, 2005.

Christian Dustmann, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. Referral-based Job Search Networks. *The Review of Economic Studies*, 83(2):514–546, 2016.

Martina Fink, Esther Kalkbrenner, Andrea Weber, and Christine Zulehner. Extracting Firm Information from Administrative Records: The ASSD Firm Panel. Technical report, NRN Working Paper: The Austrian Center for Labor Economics and the Analysis of the Welfare State, 2010.

Guido Friebel and Mariassunta Giannetti. Fighting for Talent: Risk-Taking, Corporate Volatility and Organisation Change. *The Economic Journal*, 119(540):1344–1373, 2009.

Xavier Gabaix and Augustin Landier. Why has CEO Pay Increased so Much? *The Quarterly Journal of Economics*, 123(1):49–100, 2008.

Manolis Galenianos. Learning about Match Quality and the Use of Referrals. *Review of Economic Dynamics*, 16(4):668–690, 2013.

Manolis Galenianos. Hiring through Referrals. *Journal of Economic Theory*, 152:304–323, 2014.

Albrecht Glitz. Coworker Networks in the Labour Market. *Labour Economics*, 44:218–230, 2017.

John Haltiwanger, Ron S Jarmin, and Javier Miranda. Who Creates Jobs? Small versus Large versus Young. *Review of Economics and Statistics*, 95(2):347–361, 2013.

John Haltiwanger, Henry Hyatt, and Erika McEntarfer. Cyclical Reallocation of Workers across Employers by firm size and firm wage. *National Bureau of Economic Research*, 2015.

Lena Hensvik and Oskar Nordström Skans. Social Networks, Employee Selection and Labor Market Outcomes. *Journal of Labor Economics*, 34(4):825–867, 2016.

Fredrik Heyman. Firm Size or Firm Age? The Effect on Wages using Matched Employer–employee Data. *Labour*, 21(2):237–263, 2007.

Arthur J Hosios. On the Efficiency of Matching and Related Models of Search and Unemployment. *The Review of Economic Studies*, 57(2):279–298, 1990.

Hubert Janicki, Henry Hyatt, Emin Dinlersoz, et al. Who Works for Whom? Worker Sorting in a Model of Entrepreneurship with Heterogeneous Labor Markets. *Mimeo*, 2017.

Boyan Jovanovic. Job Matching and the Theory of Turnover. *Journal of Political Economy*, 87(5, Part 1):972–990, 1979.

Boyan Jovanovic. Matching, Turnover, and Unemployment. *Journal of Political Economy*, 92(1):108–122, 1984.

Francis Kramarz and Oskar Nordström Skans. When Strong Ties are Strong: Networks and Youth Labour Market Entry. *The Review of Economic Studies*, page 1164□1200, 2014.

Fabian Lange. The Speed of Employer Learning. *Journal of Labor Economics*, 25(1):1–35, 2007.

Edward Lazear and Paul Oyer. *Handbook of Organizational Economics*. Princeton University Press, December, 2012.

Edward P Lazear. Hiring Risky Workers. *National Bureau of Economic Research*, 1995.

Edward P Lazear and Mike Gibbs. *Personnel Economics in Practice*. John Wiley & Sons, 2014.

Giuseppe Moscarini. Job Matching and the Wage Distribution. *Econometrica*, 73(2):481–516, 2005.

Giuseppe Moscarini and Fabien Postel-Vinay. The Timing of Labor Market Expansions: New Facts and a new Hypothesis. *NBER Macroeconomics Annual*, 23(1):1–52, 2008.

Giuseppe Moscarini and Fabien Postel-Vinay. Stochastic Search Equilibrium. *The Review of Economic Studies*, page rdt012, 2013.

Giuseppe Moscarini and Fabien Postel-Vinay. Did the Job Ladder Fail after the Great Recession? *Journal of Labor Economics*, 34(S1):S55–S93, 2016.

Éva Nagypál. Learning by Doing vs. Learning about Match Quality: Can we Tell them Apart? *The Review of Economic Studies*, 74(2):537–566, 2007.

Paige Ouimet and Rebecca Zarutskie. Who Works for Startups? The Relation between Firm Age, Employee Age, and Growth. *Journal of Financial Economics*, 112(3):386–407, 2014.

Paul Oyer and Scott Schaefer. Personnel Economics: Hiring and Incentives. *National Bureau of Economic Research*, 2010.

Christopher A Pissarides. Short-run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages. *The American Economic Review*, 75(4):676–690, 1985.

Michael Pries and Richard Rogerson. Hiring Policies, Labor Market Institutions, and Labor Market Flows. *Journal of Political Economy*, 113(4):811–839, 2005.

Perihan Ozge Saygin, Andrea Weber, and Michele A. Weynandt. Coworkers, Networks, and Job Search Outcomes. Technical report, IZA Discussion Paper 8174, 2014.

Scott Shane. Why Encouraging more People to Become Entrepreneurs is Bad Public Policy. *Small Business Economics*, 33(2):141–149, 2009.

Marko Terviö. The Difference that CEOs Make: An Assignment Model Approach. *The American Economic Review*, 98(3):642–668, 2008.

Josef Zweimüller, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philippe Wuellrich, Oliver Ruf, and Simon Büchi. Austrian Social Security Database. *NRN: The Austrian Center for Labor Economics and the Analysis of the Welfare State*, 2009.

3.A Additional Derivations

I show the following conditional normal distributions hold:

1.
$$m = E(y|\tilde{y}) \sim N\left(\mu, \frac{\sigma^4}{\sigma^2 + \sigma_\varepsilon^2}\right)$$

2.
$$y|\tilde{y} \sim N\left(\frac{\mu\sigma_\varepsilon^2 + \tilde{y}\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}, \frac{\sigma^2\sigma_\varepsilon^2}{\sigma^2 + \sigma_\varepsilon^2}\right)$$

Using the formula for the conditional distribution of a bivariate normal:

$$m = E(y|\tilde{y}) = \mu + \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}(\tilde{y} - \mu) = \frac{\mu\sigma_\varepsilon^2 + \sigma^2\tilde{y}}{\sigma^2 + \sigma_\varepsilon^2}$$

$$Var(m) = Var(y|\tilde{y}) = \frac{\sigma^2\sigma_\varepsilon^2}{\sigma^2 + \sigma_\varepsilon^2}$$

Hence, the distribution of $y|\tilde{y} \sim N(\frac{\mu\sigma_\varepsilon^2 + \tilde{y}\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}, \frac{\sigma^2\sigma_\varepsilon^2}{\sigma^2 + \sigma_\varepsilon^2})$.

From the equation above and the property that the random variable $\tilde{y} - \mu = y + \varepsilon - \mu$ is normal with mean zero and variance $(\sigma^2 + \sigma_\varepsilon^2)$, it follows that m (expected match quality) is itself normally distributed with mean and variance given by:

$$E(m) = \mu$$

$$Var(m) = E\left((m - \mu)^2\right) = E\left[\left(\frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}(\tilde{y} - \mu)\right)^2\right] = \frac{\sigma^4}{\sigma^2 + \sigma_\varepsilon^2}$$

Given these distributions, the mean and the variance of the mixture distributions with the mixing probability p are given by:

$$f(y) \sim pN(\mu_h, \sigma^2) + (1 - p)N(\mu_l, \sigma^2)$$

Then $E(y) = p\mu_h + (1 - p)\mu_l$ and $Var(y) = E(y^2) - E(y)^2 = p(\sigma^2 + \mu_h^2) + (1 - p)(\sigma^2 + \mu_l^2) - (p\mu_h + (1 - p)\mu_l)^2$.

Figure 3.A1 plots an example of a mixture distribution of two normal distributions. I consider the mixture of two normal distributions with different means and same standard deviation. Depending on the distance between the two means, $\mu_h - \mu_l$, the mixture distribution can become bimodal.

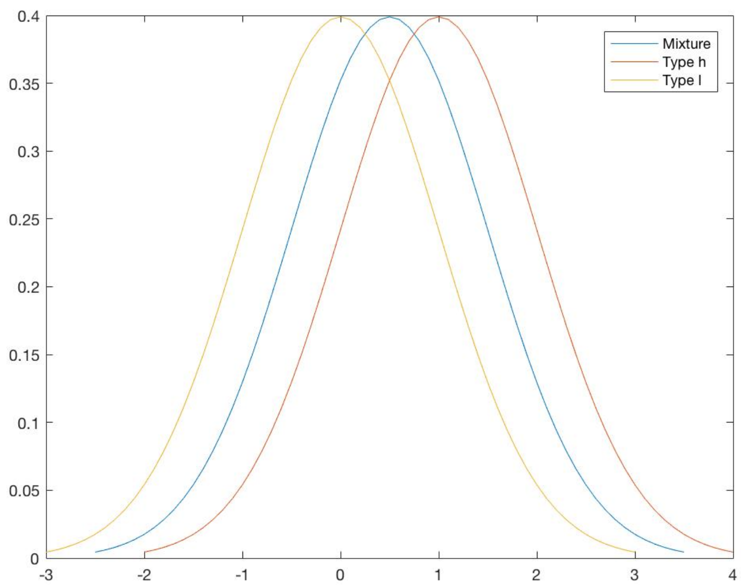


Figure 3.A1: $f(y) = 0.5N(1,1) + 0.5N(0,1)$

3.B Additional Tables and Figures

Table 3.A1 provides summary statistics for the relevant variables in the firm-level data.

Tables 3.A2 and 3.A3 have the same structure as Tables 3.1 and 3.2 but do not impose the restriction on firm survival.[12] While slightly weaker, the impact of firm age is still present and significant in the full sample. Figures 3.A2 and 3.A3 correspond to Figures 3.1 and 3.2 and show the profile of the Gini coefficient and of the number of sources over firm age when dummies for each individual year of firm age are included in the full sample.

Tables 3.A4 and 3.A5 report results for the preferred specification (column 7 of the previous tables) for 15 different sectors in the economy, i.e., Agriculture and Fishing; Construction; Education; Electricity, Gas and Water Supply; Financial Intermediation; Health and Social Work; Hotels and Restaurants; Manufacturing; Mining and Quarrying; Other community, Social and Personal service activities; Private households employing domestic staff and undifferentiated production activities of households for own use; Public administration and defense; Real Estate and Business Activities; Transport, Storage and Communication; Wholesale and Retail trade, Repair of motor vehicles and personal and household goods. The strongest effects are present in sectors where uncertainty about worker quality plays a bigger role, e.g., in Construction, Social and Personal service activities, Real Estate and Business Activities, Transport, Storage and Communication, and Wholesale and Retail trade.

[12] These results are not sensitive to changes in the age restriction to 15 or 20 years.

Table 3.A1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max	P25	P50	P75
Gini	682144	.08	.21	0	1	0	0	.07
No. source firms	682144	4.83	5.92	1	280	2	3	5
Firm age	682144	11.36	8.95	1	38	4	9	17
Firm growth	671742	1.91	21.76	-9650	1085	-1	1	4
Firm size	682144	61.32	104.84	5	1213	12	25	61
No. of beginners	682144	5.99	10.5	2	966	2	3	6
Gini NUTS-3 reg.	682144	.49	.41	0	1	.17	.35	1
Gini 4-digit ind.	682144	.19	.3	0	1	0	0	.25

Table 3.A2: Firm Age and Gini Coefficient – Full Sample

	(1) Gini	(2) Gini	(3) Gini	(4) Gini	(5) Gini	(6) Gini	(7) Gini	(8) Gini	(9) Gini
Firm age	-0.000799***	0.000530***	0.000669***	0.00133***	0.00198***	0.00165***	0.00252***	0.00321***	0.00128***
	(0.0000407)	(0.0000443)	(0.0000552)	(0.0000644)	(0.0000627)	(0.0000474)	(0.000128)	(0.000279)	(0.0000450)
Firm growth			0.0268***	0.0310***	0.0131***	0.00688***	0.00709***	0.00716***	0.00527***
			(0.000457)	(0.000480)	(0.000429)	(0.000331)	(0.000331)	(0.000332)	(0.000327)
Firm size				-0.0295***	-0.0657***	-0.0418***	-0.0432***	-0.0435***	
				(0.00136)	(0.00135)	(0.000961)	(0.000984)	(0.000991)	
No. of beginners					0.0889***	0.0495***	0.0498***	0.0499***	0.0462***
					(0.000891)	(0.000786)	(0.000787)	(0.000787)	(0.000772)
Gini NUTS-3 reg.						0.0696***	0.0696***	0.0696***	0.0703***
						(0.000838)	(0.000838)	(0.000838)	(0.000841)
Gini 4-digit ind.						0.406***	0.406***	0.406***	0.408***
						(0.00295)	(0.00295)	(0.00295)	(0.00296)
Firm age^2							-0.0000254***	-0.0000743***	
							(0.00000339)	(0.0000176)	
Firm age^3								0.000000925***	
								(0.000000323)	
Size>15									-0.0227***
									(0.00129)
Size>50									-0.0462***
									(0.00169)
Size>100									-0.0711***
									(0.00211)
Size>250									-0.103***
									(0.00289)
Size>500									-0.122***
									(0.00425)
Size>1000									-0.156***
									(0.0129)
Constant	0.0926***								
	(0.000546)								
Firm FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-3 region FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	682144	624570	340426	340426	340426	340426	340426	340426	340426
R^2	0.001	0.341	0.397	0.399	0.431	0.649	0.649	0.649	0.647

Note: Firms which, at any point in time, exceed 2 times the 99 percentile in firm size are excluded from the analysis. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

Table 3.A3: Firm Age and Number of Source Firms – Full Sample

	(1) no. source firms	(2) no. source firms	(3) no. source firms	(4) no. source firms	(5) no. source firms	(6) no. source firms	(7) no. source firms	(8) no. source firms	(9) no. source firms
Firm age	0.0187***	-0.0246***	0.00636***	-0.0466***	-0.0217***	-0.0317***	-0.0592***	-0.0837***	-0.0293***
	(0.00145)	(0.00137)	(0.00167)	(0.00160)	(0.00136)	(0.00199)	(0.00468)	(0.00872)	(0.00163)
Firm growth			1.430***	0.994***	0.275***	0.162***	0.156***	0.153***	0.206***
			(0.0176)	(0.0133)	(0.0103)	(0.0120)	(0.0120)	(0.0120)	(0.0115)
Firm size				2.761***	1.373***	1.554***	1.598***	1.609***	
				(0.0473)	(0.0391)	(0.0490)	(0.0501)	(0.0505)	
No. of beginners					3.106***	5.045***	5.035***	5.033***	4.971***
					(0.0167)	(0.0348)	(0.0347)	(0.0347)	(0.0326)
Gini NUTS-3 reg.						0.281***	0.283***	0.282***	0.169***
						(0.0153)	(0.0153)	(0.0153)	(0.0126)
Gini 4-digit ind.						-1.091***	-1.096***	-1.095***	-1.224***
						(0.0235)	(0.0235)	(0.0235)	(0.0225)
Firm age^2							0.000802***	0.00255***	
							(0.000131)	(0.000571)	
Firm age^3								-0.0000331***	
								(0.0000108)	
Size > 15									-0.155***
									(0.0180)
Size > 50									0.275***
									(0.0387)
Size > 100									1.999***
									(0.0713)
Size > 250									6.519***
									(0.187)
Size > 500									13.89***
									(0.472)
Size > 1000									24.15***
									(1.813)
Constant	2.913***								
	(0.0157)								
Firm FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-3 region FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1227652	1166554	577441	577441	577441	340426	340426	340426	340426
R^2	0.001	0.661	0.754	0.774	0.839	0.844	0.844	0.844	0.857

Note: Firms which, at any point in time, exceed 2 times the 99 percentile in firm size are excluded from the analysis. Standard errors are clustered at the firm level. Significance levels are given by * < 0.1, ** < 0.05, and *** < 0.01.

Table 3.A4: Firm Age and Gini Coefficient, by Sector

	(1) Gini	(2) Gini	(3) Gini	(4) Gini	(5) Gini	(6) Gini	(7) Gini	(8) Gini	(9) Gini	(10) Gini	(11) Gini	(12) Gini	(13) Gini	(14) Gini	(15) Gini
Firm age	0.00219	0.00255***	-0.00118	0.000700	0.000660	0.00210***	0.00121***	0.00317***	0.00202	0.00195***	-0.00943	0.00271***	0.00298***	0.00246***	0.00290***
	(0.00157)	(0.000387)	(0.00138)	(0.00176)	(0.000746)	(0.000664)	(0.000428)	(0.000237)	(0.00143)	(0.000637)	(0.00681)	(0.000693)	(0.000479)	(0.000467)	(0.000268)
Firm_age2	-0.0000346	-0.0000195*	0.0000429	0.0000228	0.0000212	-0.0000382**	-0.0000279**	-0.0000266***	-0.0000105	-0.0000259*	0.0000594	-0.0000137	-0.0000444***	-0.0000457***	-0.0000327***
	(0.0000437)	(0.0000107)	(0.0000381)	(0.0000482)	(0.0000167)	(0.0000162)	(0.0000117)	(0.00000620)	(0.0000364)	(0.0000156)	(0.000158)	(0.0000193)	(0.0000132)	(0.0000121)	(0.00000720)
Firm growth	0.00742*	0.00316***	0.00234	0.00536	0.00690***	0.00513***	0.0109***	0.00420***	0.00163	0.00685***	0.0232	0.00361**	0.00724***	0.0155***	0.00800***
	(0.00430)	(0.00104)	(0.00365)	(0.00392)	(0.00157)	(0.00158)	(0.00132)	(0.000604)	(0.00407)	(0.00164)	(0.0277)	(0.00150)	(0.00109)	(0.00136)	(0.000680)
Firm size	-0.0159	-0.0481***	-0.0393***	-0.0296**	-0.0375***	-0.0417***	-0.0287***	-0.0397***	-0.0374**	-0.0405***	0.185**	-0.0537***	-0.0514***	-0.0467***	-0.0359***
	(0.0120)	(0.00322)	(0.00774)	(0.0131)	(0.00728)	(0.00498)	(0.00318)	(0.00190)	(0.0174)	(0.00489)	(0.0834)	(0.00526)	(0.00295)	(0.00376)	(0.00201)
No. of beginners	0.0548***	0.0670***	0.0412***	0.0338***	0.0490***	0.0424***	0.0446***	0.0359***	0.0378***	0.0526***	-0.0367	0.0291***	0.0539***	0.0547***	0.0428***
	(0.0111)	(0.00225)	(0.00963)	(0.0126)	(0.00427)	(0.00470)	(0.00267)	(0.00162)	(0.00959)	(0.00397)	(0.0427)	(0.00434)	(0.00265)	(0.00269)	(0.00171)
Gini NUTS-3 reg.	0.117***	0.0904***	0.0540***	0.0575***	0.0772***	0.0616***	0.0711***	0.0502***	0.0662***	0.0774***	0.0242	0.0690***	0.0626***	0.0847***	0.0585***
	(0.0103)	(0.00260)	(0.0105)	(0.0148)	(0.00505)	(0.00491)	(0.00322)	(0.00163)	(0.00967)	(0.00439)	(0.0298)	(0.00485)	(0.00267)	(0.00322)	(0.00165)
Gini 4-digit ind.	0.543***	0.374***	0.576***	0.634***	0.336***	0.446***	0.206***	0.497***	0.459***	0.395***	0.730***	0.595***	0.474***	0.272***	0.484***
	(0.0276)	(0.00700)	(0.0389)	(0.0512)	(0.0154)	(0.0187)	(0.00751)	(0.00765)	(0.0478)	(0.0144)	(0.121)	(0.0172)	(0.00989)	(0.00815)	(0.00725)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-3 region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3090	45033	2588	1403	10852	10705	24995	70594	1967	13905	100	11279	33118	27048	69460
R^2	0.803	0.636	0.758	0.757	0.610	0.652	0.514	0.681	0.687	0.637	0.843	0.767	0.696	0.553	0.689

Note: Firms which, at any point in time, exceed 2 times the 99 percentile in firm size are excluded from the analysis. Standard errors are clustered at the firm level. Significance levels are given by $^{*} < 0.1$, $^{**} < 0.05$, and $^{***} < 0.01$.

Table 3.A5: Firm Age and Number of Source Firms, by Sector

	(1) no. hiring firms	(2) no. hiring firms	(3) no. hiring firms	(4) no. hiring firms	(5) no. hiring firms	(6) no. hiring firms	(7) no. hiring firms	(8) no. hiring firms	(9) no. hiring firms	(10) no. hiring firms	(11) no. hiring firms	(12) no. hiring firms	(13) no. hiring firms	(14) no. hiring firms	(15) no. hiring firms
firm age	-0.0473**	-0.114***	0.0278	0.0816	-0.0401**	-0.0439	-0.0248**	-0.0493***	-0.0750**	-0.0314*	0.00952	-0.00839	-0.0923***	-0.0804***	-0.0635***
	(0.0214)	(0.0124)	(0.0480)	(0.0529)	(0.0195)	(0.0310)	(0.0106)	(0.00972)	(0.0323)	(0.0186)	(0.0545)	(0.0315)	(0.0217)	(0.0248)	(0.00941)
firm_age2	0.00119*	0.00178***	-0.000173	-0.00308**	0.000569	0.00157**	0.000323	0.000122	0.00164*	0.000403	-0.000222	-0.000878	0.00147**	0.00176***	0.00126***
	(0.000615)	(0.000370)	(0.00150)	(0.00151)	(0.000421)	(0.000772)	(0.000293)	(0.000267)	(0.000936)	(0.000517)	(0.00104)	(0.000871)	(0.000590)	(0.000587)	(0.000282)
firm growth	-0.00571	0.143***	0.0634	0.0820	0.0313	0.226***	0.0805**	0.337***	0.143	0.0629	-0.111	0.266***	0.0805*	-0.0696	0.125***
	(0.0526)	(0.0311)	(0.110)	(0.161)	(0.0410)	(0.0586)	(0.0349)	(0.0249)	(0.101)	(0.0492)	(0.102)	(0.0841)	(0.0449)	(0.0603)	(0.0249)
firm size	0.561***	2.022***	1.536***	0.889**	0.906***	1.398***	0.682***	1.394***	1.199***	1.351***	0.475	1.989***	2.490***	1.611***	1.459***
	(0.163)	(0.144)	(0.451)	(0.372)	(0.174)	(0.314)	(0.117)	(0.0909)	(0.393)	(0.385)	(0.372)	(0.326)	(0.182)	(0.265)	(0.0930)
no. of beginners	2.498***	4.499***	3.950***	4.832***	5.358***	5.711***	4.760***	5.740***	4.734***	4.765***	2.843***	5.361***	5.049***	5.120***	4.683***
	(0.172)	(0.0808)	(0.268)	(0.450)	(0.177)	(0.169)	(0.0956)	(0.0764)	(0.260)	(0.233)	(0.556)	(0.296)	(0.106)	(0.155)	(0.0658)
Gini NUTS-3 reg.	-0.323***	0.0999***	0.219**	0.181	0.249***	0.577***	0.164***	0.551***	0.215**	0.144	-0.0671	0.318***	0.407***	0.261***	0.243***
	(0.0645)	(0.0369)	(0.111)	(0.247)	(0.0782)	(0.0892)	(0.0376)	(0.0387)	(0.101)	(0.0923)	(0.141)	(0.0962)	(0.0562)	(0.0624)	(0.0305)
Gini 4-digit ind.	-1.088***	-0.694***	-2.075***	-3.080***	-0.980***	-1.119***	-0.0523	-2.049***	-1.348***	-0.903***	-0.824***	-1.842***	-1.365***	-0.546***	-1.435***
	(0.0979)	(0.0444)	(0.245)	(0.587)	(0.109)	(0.142)	(0.0422)	(0.0820)	(0.285)	(0.121)	(0.298)	(0.174)	(0.0851)	(0.0685)	(0.0517)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-3 region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3090	45033	2588	1403	10852	10705	24995	70594	1967	13905	100	11279	33118	27048	69460
R^2	0.755	0.845	0.889	0.827	0.859	0.877	0.870	0.828	0.840	0.802	0.879	0.824	0.867	0.828	0.838

Note: Firms which, at any point in time, exceed 2 times the 99 percentile in firm size are excluded from the analysis. Standard errors are clustered at the firm level. Significance levels are given by $^{*} < 0.1$, $^{**} < 0.05$, and $^{***} < 0.01$.

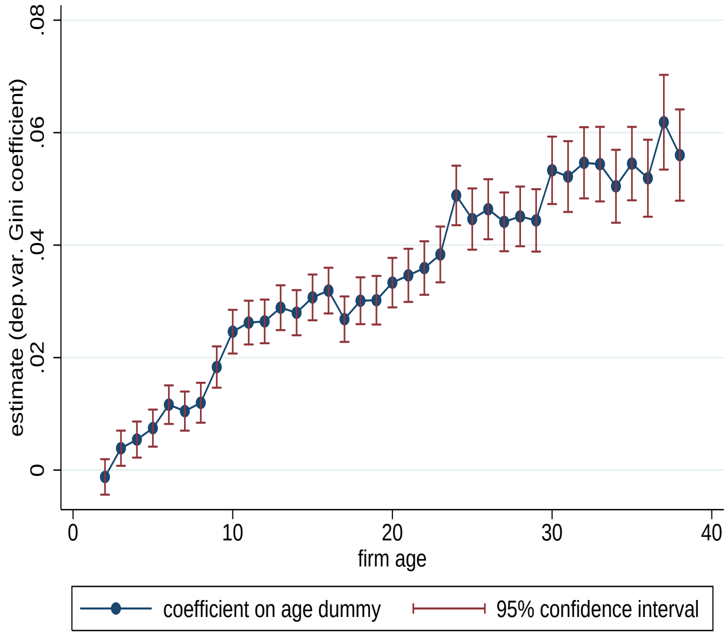


Figure 3.A2: Gini Index and Age (Estimated Coefficients on Age Dummies)

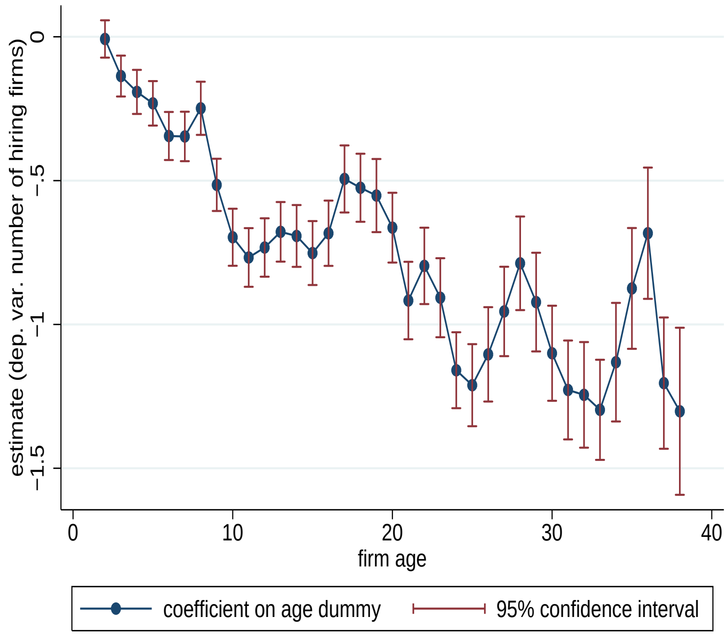


Figure 3.A3: Number of Source Firms and Age (Estimated Coefficients on Age Dummies)

Chapter 4

Using the Weighted Bootstrap to Account for Clustered Standard Errors

with Bernd Fitzenberger

4.1 Introduction

Many econometric applications involve data sets with clustered observations while the number of clusters is often limited by geographic and administrative reasons or budget considerations. We examine the performance of bootstrap methods that explicitly allow for models which include cluster-invariant binary variables in a small-sample context.

In cases with a small number of clusters, conventional adjustments of standard errors in linear estimation models are biased. Cameron, Gelbach, and Miller (2008) show that in case of few clusters the use of bootstrap methods can provide asymptotic refinement on approximations from asymptotic theory. However, a variety of numerical problems arises with the use of these methods. First, the most precise strategies are those that bootstrap studentized statistics. As a drawback, they do not allow to estimate standard errors directly. Second and at the main focus of this paper, with only few clusters conventional bootstrap methods fail to estimate models with cluster-invariant binary variables (see Cameron and Miller, 2015).[1] This extends to the use of non-linear models like probit. To address these problems we introduce and evaluate the (more generalized) weighted bootstrap (Barbe and Bertail, 1995; Mason and Newton, 1992) which generates resamples by repeatedly attaching random weights to each cluster in the data set. We provide Monte Carlo evidence for small but persistent asymptotic refinement of a uniformly weighted bootstrap algorithm over conventional bootstrap methods.

[1] Moreover, Mackinnon and Webb (2016) show that small G-problems are amplified if cluster sizes are asymmetric.

4.2 Clustered Standard Errors

Assume a simple random effects model for $i \in \{1, \ldots, N\}$ individuals in $g \in \{1, \ldots, G(< N)\}$ clusters,

$$y_{ig} = \mathrm{x}'_{ig}\beta + u_{ig}, \tag{4.1}$$

where the error term $u_{ig} = \alpha_g + \varepsilon_{ig}$ is composed of a group component α_g and an idiosyncratic term ε_{ig}. For simplicity assume balanced clusters, i.e., $N_g = \bar{N}_g \; \forall g$. Standard errors can be consistently estimated by (square roots of diagonal entries of)

$$\widehat{Avar}_{CR}(\hat{\beta}) = (\mathrm{X}'\mathrm{X})^{-1} \left(\sum_{g=1}^{G} \mathrm{X}_g \tilde{u}_g \tilde{u}'_g \mathrm{X}'_g \right) (\mathrm{X}'\mathrm{X})^{-1}. \tag{4.2}$$

Inference in a linear regression model with grouped data however relies on the crucial assumption that G tends to infinity. Explicitly calculating the analytical standard errors in small samples with complex dependence structures is a very complicated process and has no practical relevance. Using bootstrap methods therefore is an attractive approach to approximate the unknown distribution of regression coefficients with only few clusters. Efron's (1979) classical bootstrap algorithm can be extended to the case of clustered data, a procedure often called the pairs block bootstrap. The dependence structure in each bootstrap resample should mirror the original sample as closely as possible. Therefore, entire clusters are drawn in each of B resamples to preserve the within-cluster correlation structure (Cameron et al., 2008). The bootstrap estimator that approximates the finite sample cdf of some statistic T_G, $F_G(x) = Pr[T_G \leq x]$, is obtained from the empirical distribution function of the data, $\hat{F}_G(x)$. This ECDF in turn is approximated by

$$\widehat{F}_G(x) = B^{-1} \sum_{b=1}^{B} 1\{T^*_{G,b} \leq x\}. \tag{4.3}$$

where $T^*_{G,b}$ denotes the statistic of interest calculated from resample $b \in \{1, \ldots, B\}$. In many instances, however, the conventional scheme of random resampling with replacement leads to numerical problems. In particular, caused by the use of cluster-invariant binary explanatory variables resamples with perfect collinearity of the regressors are likely to occur if the number of groups is small.[2] A weighted bootstrap algorithm avoids these problems by employing a more generalized resampling scheme. In particular, each cluster from the original sample appears in each bootstrap resample while the random variation in the procedure stems from randomly generated weights that are attached to each cluster. The weighted bootstrap

[2] A similar problem arises for a standard pairwise bootstrap in a probit or logit model when the resample accidentally contains only observations with the dependent variable being equal to zero (or one) for all observations in certain cells (Consequently, the probit would try to predict an exact zero [or an exact one] for such observations).

distribution thus becomes

$$\widehat{F}_{\mathcal{W},G}(x) = \sum_{b=1}^{B} W_{G,b} 1\{T^*_{G,b} \leq x\}. \tag{4.4}$$

where $W_{G,b}$ denotes cluster specific weights that are generated drawn from a distribution $\mathcal{W}_G$.[3] However, there are some conditions which have to be satisfied for consistency of the weighted bootstrap due to Mason and Newton (1992) and Barbe and Bertail (1995). In particular,

$$W_{g,G} \geq 0, \qquad g = 1, \dots, G, \quad G \geq 1, \tag{4.5a}$$

$$G^{-1} \sum_{g=1}^{G} W_{g,G} = 1, \tag{4.5b}$$

$$G^{-1} \sum_{g=1}^{G} (W_{g,G} - 1/G)^2 \xrightarrow{p} c \quad \text{as } G \to \infty \quad \text{for some } c > 0, \tag{4.5c}$$

$$\text{and } \mathcal{W}_G \text{ and the } W_{g,G}\text{'s are independent of the sample.} \tag{4.5d}$$

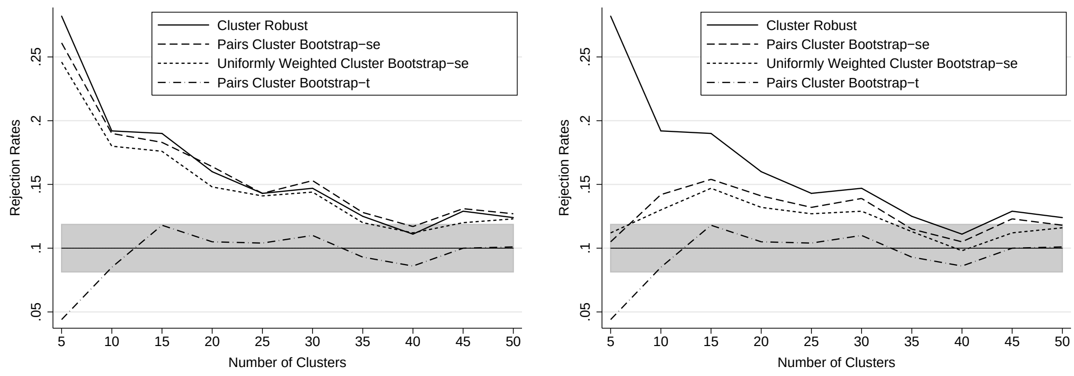
Potential distributions for the weights include the exponential distribution with $\lambda = 1$ as discussed in a related context by Chernozukov, Fernández-Val, and Kowalski (2015) as well as the uniform distribution on $(1-a, 1+a)$. Importantly, the bootstrap estimate of the standard error has to be adjusted for the variance of the weights $V[W_G]$.

The following scheme spells out in detail how the weighted bootstrap algorithm is conducted (using the uniform distribution):

1. Do B iterations of this step. On the b^{th} iteration:
   (a) Randomly draw G weights from the $\mathcal{U}$niform$(1 - a, 1 + a)$ distribution.
   (b) Form a resample of G clusters $\{(\mathrm{y}^*_1, \mathrm{X}^*_1), \dots, (\mathrm{y}^*_G, \mathrm{X}^*_G)\}$ by multiplying the original sample with the square roots of the random weights $W_1, \dots, W_G$.
   (c) Calculate the (weighted) OLS estimate $\hat{\beta}^*_{j,b}$ by regressing y^* on X^*.
2. Reject $H_0 : \beta_j = \beta^0_j$ at level α if and only if $|w| > z_{1-\alpha/2}$, where $w = (\hat{\beta}_j - \beta^0_j)/s_{\hat{\beta}_{j,B}}$. The bootstrap estimate of the standard error is adjusted for the variance of the weights $Var[W_G]$ as follows:

$$s_{\hat{\beta}_{j,B}} = \left(\left(\tfrac{1}{12}(2a)^2 \right)^{-1} \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\beta}^*_{j,b} - \overline{\hat{\beta}^*_j})^2 \right)^{1/2}.$$

[3]Note that the pairs cluster bootstrap is a special case with $\mathcal{W}_G = G^{-1}\mathcal{M}ultinomial(G; 1/G, \dots, 1/G)$.



(a) Critical Values from the Standard Normal Distribution

(b) Critical Values from the t-Distribution

Figure 4.1: Rejection Rates in the Homoscedastic Random Effects Model

4.3 Monte-Carlo Simulations

Cameron et al. (2008) provide a comparison of various methods that account for cluster effects concentrating on bootstrap algorithms. Their main finding is that methods which bootstrap studentized statistics (such as the t-statistic) provide an asymptotic refinement that improves on convergences rates from asymptotic theory even with a small number of clusters. However, these methods do not allow to compute standard errors directly. Moreover, as spelled out in section 4.2 most methods work poorly if the model includes cluster-invariant binary variables. Using the same model and methodology as Cameron et al. (2008) we extend their analysis by considering the weighted bootstrap algorithm presented above.

In particular, we use various data generating processes to generate the data in $R = 1000$ Monte-Carlo simulations and report the fraction of simulations that rejects $H_0 : \beta_j = \beta_j^0$ because w falls outside the 90% confidence interval. The nominal rejection rate hence amounts to $\alpha = 10\%$ while the simulation error becomes $s_{\hat{\alpha}} = \sqrt{\hat{\alpha}(1-\hat{\alpha})/(R-1)}$.

The main specification with homoscedastic errors reads

$$y_{ig} = \beta_0 + \beta_1 x_{ig} + u_{ig} = \beta_0 + \beta_1(z_g + z_{ig}) + (\alpha_g + \varepsilon_{ig}), \tag{4.6}$$

where z_g, z_{ig}, ε_g, and ε_{ig} each are independent draws from $\mathcal{N}(0, 0.1)$. The (true) coefficients are $\beta_0 = 0$ and $\beta_1 = 1$.[4] The analysis is conducted for various group sizes, $G = 5, 10, 15, \ldots, 50$, while (in the baseline configuration) the number of observations per group is held constant at $N_g = 30$. Our simulation results are summarized in figure 4.1 while detailed numbers can be found in table 4.1. The results in panel (a) confirm the findings in Cameron et al. (2008).

In addition, the weighted cluster bootstrap based on uniformly distributed weights achieves rejection rates that are minimally but persistently smaller than the pairs cluster bootstrap. Simple t-tests for dependent samples indicate that this difference is significant on the 5% con-

[4] As a result, the within-group correlations $\varrho_u = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\varepsilon^2)$ and equivalently ϱ_x both are equal to 0.5.

Table 4.1: Rejection Rates from 1000 Monte-Carlo Simulations of the Homoscedastic Random Effects Model

	Number of Groups G									
Method	5	10	15	20	25	30	35	40	45	50
cluster robust standard errors	0.282	0.192	0.190	0.160	0.143	0.147	0.125	0.111	0.129	0.124
	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)
pairs cluster bootstrap-se	0.261	0.190	0.183	0.164	0.143	0.153	0.128	0.117	0.131	0.127
	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)
pairs cluster bootstrap-se (T_{G-2})	0.105	0.142	0.154	0.141	0.132	0.139	0.115	0.105	0.123	0.118
	(0.009)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
uniformly weighted cluster bootstrap-se	0.246	0.180	0.176	0.148	0.141	0.144	0.120	0.112	0.120	0.123
	(0.013)	(0.013)	(0.011)	(0.012)	(0.012)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
uniformly weighted cluster bootstrap-se (T_{G-2})	0.112	0.130	0.147	0.132	0.127	0.129	0.113	0.098	0.112	0.116
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)
exponentially weighted cluster bootstrap-se	0.385	0.262	0.240	0.204	0.170	0.175	0.144	0.129	0.150	0.145
	(0.016)	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
exponentially weighted cluster bootstrap-se (T_{G-2})	0.214	0.208	0.211	0.179	0.154	0.164	0.136	0.120	0.140	0.137
	(0.016)	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
pairs cluster bootstrap-t	0.044	0.085	0.118	0.105	0.104	0.110	0.093	0.086	0.100	0.101
	(0.007)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.009)

Table 4.2: Test Statistics for Comparison of Rejection Rates of Uniformly Weighted Cluster Bootstrap vs. Conventional Methods - Homoscedastic Model

Number of groups G	Test statistics for H_0: unif. WCB $\geq$		
	CRVE	pairs cluster bootstrap-se	pairs cluster bootstrap-se (T_{G-2})
5	6.11***	2.54***	-1.22
10	3.01***	2.14**	3.01***
15	3.52***	1.53	1.81*
20	2.83***	4.03***	2.50***
25	0.63	0.58	1.29
30	0.83	2.19**	2.36***
35	1.39	2.53***	0.63
40	-0.38	1.39	1.81*
45	2.72***	3.07***	3.34***
50	0.38	1.07	0.53

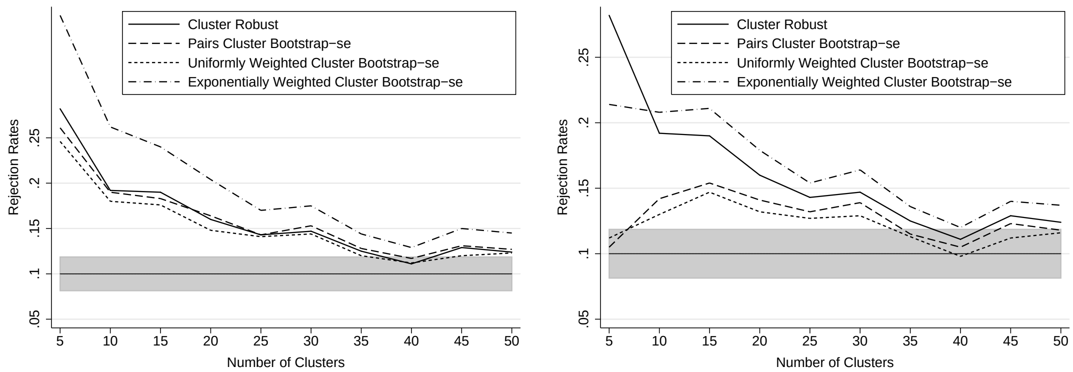
Note: ***, **, and * denote rejections on the 1, 5, and 10% level, respectively.

fidence level for almost all cluster sizes (see test statistics in table 4.2). In panel (b), critical values from the t-distribution with $G - 2$ degrees of freedom are used for the bootstrap-se methods and lead to largely improved rejection rates.

Several weighting schemes different from the uniform distribution cannot reproduce the favorable results of the uniformly weighted cluster bootstrap. Figure 4.2 shows that in our setup using the exponential distribution as suggested by Chernozukov et al. (2015) leads to much higher rejection rates for all cluster sizes with both sets of critical values.

Particular interest is in the performance of the weighted cluster bootstrap in a model with a binary regressor where conventional bootstrap methods have no bite. Therefore, we modify the baseline specification to a variant with a cluster-invariant regressor, $x_{ig} = x_g = 1\{g \geq 0.8\,G\}$, i.e., a model with an indicator function that equals one for 20% of the clusters. Since there is no natural ordering of the clusters this can be interpreted as a random treatment dummy variable on the cluster level with $P(X_g = 1) = 0.2$. Results are shown in figure 4.3 (and table 4.3).

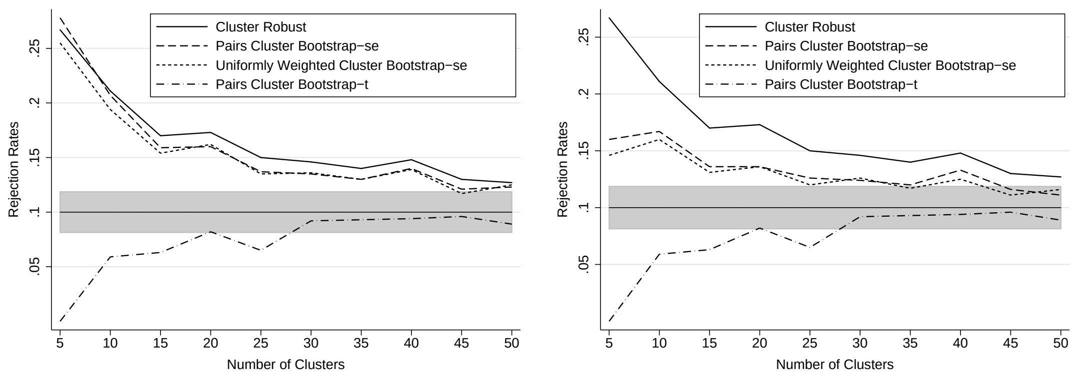
The rejection rates for the pairs cluster bootstrap cannot be interpreted due to the fact that (in particular for small G) many bootstrap replications sample only clusters with $x_g = 0$ (or 1). In these replications, the statistical package estimates a perfect fit for the treatment indicator and hence "zero" coefficients. Since in (re-)samples with variation in x_g the algorithm consistently estimates the true $\beta_1 = 1$ this vastly increases variation in $\hat{\beta}^*_{j,b}$ and hence artificially decreases the test statistic w. This mechanism generates under-rejection rendering the



(a) Critical Values from the Standard Normal Distribution

(b) Critical Values from the *t*-Distribution

Figure 4.2: Rejection Rates in the Homoscedastic Random Effects Model



(a) Critical Values from the Standard Normal Distribution

(b) Critical Values from the *t*-Distribution

Figure 4.3: Rejection Rates in the Model with Binary Regressor

Table 4.3: Rejection Rates from 1000 Monte-Carlo Simulations of the Random Effects Model with Binary Regressor

	Number of Groups G									
Method	5	10	15	20	25	30	35	40	45	50
cluster robust standard errors	0.267	0.211	0.170	0.173	0.150	0.146	0.140	0.148	0.130	0.127
	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)
pairs cluster bootstrap-se	0.278	0.207	0.159	0.160	0.137	0.135	0.130	0.140	0.121	0.123
	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)
pairs cluster bootstrap-se (T_{G-2})	0.160	0.167	0.136	0.136	0.126	0.124	0.120	0.133	0.116	0.111
	(0.009)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
uniformly weighted cluster bootstrap-se	0.255	0.194	0.154	0.162	0.135	0.136	0.130	0.139	0.117	0.125
	(0.013)	(0.013)	(0.011)	(0.012)	(0.012)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
uniformly weighted cluster bootstrap-se (T_{G-2})	0.146	0.160	0.131	0.136	0.120	0.126	0.117	0.125	0.111	0.116
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)
exponentially weighted cluster bootstrap-se	0.375	0.264	0.225	0.113	0.178	0.178	0.160	0.168	0.146	0.149
	(0.016)	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
exponentially weighted cluster bootstrap-se (T_{G-2})	0.244	0.220	0.193	0.186	0.160	0.157	0.146	0.160	0.137	0.137
	(0.016)	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
pairs cluster bootstrap-t	0.000	0.059	0.063	0.82	0.065	0.092	0.093	0.094	0.096	0.089
	(0.007)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.009)

Table 4.4: Test Statistics for Comparison of Rejection Rates of Uniformly Weighted Cluster Bootstrap vs. Conventional Methods - Model with Binary Regressor

Number of groups G	Test statistics for H_0: unif. WCB $\geq$		
	CRVE	pairs cluster bootstrap-se	pairs cluster bootstrap-se (T_{G-2})
5	3.22***	4.47***	3.77***
10	4.16***	3.00***	1.95*
15	4.03***	1.29	1.15
20	2.83***	-0.43	0
25	3.66***	0.53	1.61
30	3.18***	-0.28	-0.58
35	2.51***	0.00	0.83
40	2.72***	0.25	2.14**
45	3.63***	1.07	1.29
50	0.71	-0.47	-1.29

Note: ***, **, and * denote rejections on the 1, 5, and 10% level, respectively.

results for the pairs cluster bootstrap-se misleading.[5] The pairs cluster bootstrap-t displays substantial underrejection for small G due to a different mechanism. Here, perfect fit of estimates induces zero standard errors in a resample and thus generates missing values for the test statistic being bootstrapped. As laid out in Cameron and Miller (2015) and confirmed in both panels of figure 4.3 this effect leads to underrejection. The weighted cluster bootstrap in contrast is not subject to these limitations and delivers similar results as before. Comparison to rejection rates from cluster-robust variance estimates (column CRVE in table 4.4) indicates a clear improvement particularly for small G while the other columns should be interpreted with caution due to the problems of the pairs cluster bootstrap with binary data.

In addition, we perform multiple sensitivity checks varying the parameters of the model. The performance of the weighted cluster bootstrap with uniform weights is remarkably stable. Introducing heteroscedasticity invalidates residual bootstraps but does not affect the relative performance of the weighted cluster boostrap. If the estimated model is misspecified because a quadratic term of x_{ig}^2 is omitted differences in favor of the uniformly weighted cluster bootstrap become slightly larger.

[5]We cannot give formal evidence for the size of this effect. In the smallest configuration with 5 clusters, however, the probability of drawing a resample that contains only zeros amounts to $(4/5)^5 = 0.32768$ indicating that this indeed leads to problems with the interpretation of the results for the pairs cluster bootstrap-se in figure 4.3.

References

P. Barbe and P. Bertail. *The Weighted Bootstrap*. Springer, New York, 1995.

A. Colin Cameron and Douglas L. Miller. A Practitioner□s Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372, March 2015.

A Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, 90(3):414–427, 2008.

V. Chernozukov, Iván Fernández-Val, and Amanda Kowalski. Quantile Regression with Censoring and Endogeneity. *Journal of Econometrics*, 186:201–221, May 2015.

B. Efron. Bootstrapping Methods: Another Look at the Jackknife. *Annals of Statistics*, 7: 1–26, 1979.

James G. Mackinnon and Matthew D. Webb. Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics*, 2016. jae.2508.

D. M. Mason and M. A. Newton. A Rank Statistics Approach to the Consistency of a General Bootstrap. *The Annals of Statistics*, 20(3):1611–1624, 1992.

Conclusion

In this thesis, I provide ample evidence for the importance of job mobility and network structures for the functioning of labor markets. Patterns of job mobility can be used to explain the formation of separate labor markets within countries, to understand information transmission between individuals, or to understand mechanisms of matching and hiring strategies. Repeated interaction between individuals or firms can lead to the formation of stable networks and groups which are important factors in diminishing informational uncertainties. I propose novel econometric methods that allow to explicitly incorporate interactions and dependency structures using techniques from network analysis and machine learning.

The research in this thesis lays ground for a range of important follow-up questions that I aim to answer in future projects. First, it is interesting to examine and understand the hierarchical structure that underlies many transitions between jobs. Individuals aim to advance to better firms, occupations, or tasks. The network methodology proposed in this thesis allows to empirically analyze the concept of job ladders and might be helpful in answering important questions regarding the functioning of the labor market. Second, while the results in this thesis indicate that job mobility is an important driver of information flow and labor market efficiency, an open economic question pertains to the consequences of obstacles to worker flows. In future research, I aim to study the consequences of particular historic events that resulted in impediments to worker mobility. Two episodes of German history – policy differences between occupation zones after World War II and the subsequent division of the country in two parts – create unique natural experiments that can be exploited in order to identify the consequences of impediments to mobility for economic growth, labor market integration, and other socio-economic outcomes.

Bibliography

Bibliography

John M. Abowd, Francis Kramarz, and David N. Margolis. High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333, 1999.

John M. Abowd, Robert Creecy, and Francis Kramarz. Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Unpublished working paper, Cornell University Department of Economics, March 2002.

Daron Acemoglu and Joshua Angrist. How Large are Human-capital Externalities? Evidence from Compulsory-schooling Laws. In *NBER Macroeconomics Annual 2000, Volume 15*, pages 9–74. MIT Press, 2000.

Christopher Aicher, Abigail Z. Jacobs, and Aaron Clauset. Learning Latent Block Structure in Weighted Networks. *Journal of Complex Networks*, 3(2):221–248, 2015.

Edoardo M Airoldi, David M Blei, Stephen E Fienberg, and Eric P Xing. Mixed Membership Stochastic Blockmodels. In *Advances in Neural Information Processing Systems*, pages 33–40, 2009.

Joseph G Altonji and Robert A Shakotko. Do Wages Rise with Job Seniority? *The Review of Economic Studies*, 54(3):437–459, 1987.

Fredrik Andersson, Matthew Freedman, John Haltiwanger, Julia Lane, and Kathryn Shaw. Reaching for the Stars: Who Pays for Talent in Innovative Industries? *The Economic Journal*, 119(538):F308–F332, 2009.

David H. Autor, David Dorn, and Gordon H. Hanson. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6):2121–2168, 2013.

David H. Autor, David Dorn, Gordon H. Hanson, and Jae Song. Trade Adjustment: Worker Level Evidence. *Quarterly Journal of Economics*, 129(4), 2014.

Pierre Bachas and Mauricio Soto. Not(ch) Your Average Tax System: Corporate Taxation Under Weak Enforcement. Technical report, UC Berkeley, 2015.

Jesper Bagger, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics. *The American Economic Review*, 104(6):1551–1596, 2014.

Ragnhild Balsvik. Is Labor Mobility a Channel for Spillovers from Multinationals? Evidence from Norwegian Manufacturing. *The Review of Economics and Statistics*, 93(1):285–297, 2011.

P. Barbe and P. Bertail. *The Weighted Bootstrap*. Springer, New York, 1995.

Spencer Bastani and Håkan Selin. Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics*, 109:36–49, 2014.

Patrick Bayer, Stephen L Ross, and Giorgio Topa. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116 (6):1150–1196, 2008.

Thomas Behrends. Recruitment Practices in Small and Medium Size Enterprises. An Empirical Study among Knowledge-intensive Professional Service Firms. *Management Revue*, pages 55–74, 2007.

Annette Bergmann and Antje Mertens. Job Stability Trends, Lay-offs, and Transitions to Unemployment in West Germany. *Labour*, 25(4):421–446, 2011.

Timothy J Besley and Torsten Persson. Taxation and Development. *Handbook of Public Economics*, 5:51–110, 2013.

Michael Best, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan. *Journal of Political Economy*, 123(6):1311–1355, 2015.

Peter J Bickel and Aiyou Chen. A Nonparametric View of Network Models and Newman–Girvan and other Modularities. *Proceedings of the National Academy of Sciences*, 106(50): 21068–21073, 2009.

Melissa Bjelland, Bruce Fallick, John Haltiwanger, and Erika McEntarfer. Employer-to-Employer Flows in the United States: Estimates Using Linked Employer-Employee Data. *Journal of Business & Economic Statistics*, 29(4):493–505, 2011.

Olivier Blanchard and Larry Katz. Regional Evolutions. *Brookings Papers on Economic Activity*, 1:1–75, 1992.

Stéphane Bonhomme and Elena Manresa. Grouped Patterns of Heterogeneity in Panel Data. *Econometrica*, 83(3):1147–1184, 2015.

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. A Distributional Framework for Matched Employer Employee Data. Unpublished manuscript, University of Chicago, 2016.

Katarına Borovicková. Job Flows, Worker Flows and Labor Market Policies. *Mimeo*, 2016.

John Bound and Harry Holzer. Demand Shifts, Population Adjustments and Labor Market Outcomes during the 1980s. *Journal of Labor Economics*, 18:20–54, 2000.

Udo Brixy, Susanne Kohaut, and Claus Schnabel. Do Newly Founded Firms Pay Lower Wages? First Evidence from Germany. *Small Business Economics*, 29(1):161–171, 2007.

Peter J Brockwell and Richard A Davis. *Introduction to Time Series and Forecasting*. Springer, 2016.

Charles Brown and James Medoff. The Employer Size-wage Effect. *Journal of Political Economy*, 97(5):1027–1059, 1989.

Charles Brown and James L Medoff. Firm Age and Wages. *Journal of Labor Economics*, 21 (3):677–697, 2003.

Kenneth Burdett and Dale T. Mortensen. Wage Differentials, Employer Size, and Unemployment. *International Economic Review*, 39(2):257–273, 1998.

Simon Burgess, Julia Lane, and David Stevens. Hiring Risky Workers: Some Evidence. *Journal of Economics & Management Strategy*, 7(4):669–676, 1998.

Matias Busso, Jesse Gregory, and Patrick Kline. Assessing the Incidence and Efficiency of a Prominent Place Based Policy. *American Economic Review*, 103(2):897–947, 2013.

A. Colin Cameron and Douglas L. Miller. A Practitioner□s Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372, March 2015.

A Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, 90(3):414–427, 2008.

Carla Canelas. Minimum Wage and Informality in Ecuador. Technical report, UNU-WIDER Working Paper, 2014.

David Card, Jörg Heining, and Patrick Kline. Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128, August 2013.

Paul Carrillo, M. Shahe Emran, and Anita Rivadeneira. Do Cheaters Bunch Together? Profit Taxes, Wothholding Rates and Tax Evasion. Technical report, Working Paper, 2012.

Paul Carrillo, Dina Pomeranz, and Monica Singhal. Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement. *American Economic Journal: Applied Economics*, 9(2): 144–64, April 2017.

Vasco M. Carvalho. From Micro to Macro via Production Networks. *Journal of Economic Perspectives*, 28(4):23–48, 2014.

Giacinta Cestone, Chiara Fumagalli, Francis Kramarz, and Giovanni Pica. Insurance Between Firms: The Role of Internal Labor Markets. Technical Report DP11336, CEPR, June 2016.

Thomas Chaney. The Network Structure of International Trade. *American Economic Review*, 104:3600–3634, 2014.

V. Chernozukov, Iván Fernández-Val, and Amanda Kowalski. Quantile Regression with Censoring and Endogeneity. *Journal of Econometrics*, 186:201–221, May 2015.

Raj Chetty and Emmanuel Saez. Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients. *American Economic Journal: Applied Economics*, 5(1):1–31, 2013.

Raj Chetty, John Friedman, Tore Olsen, and Luigi Pistaferri. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *Quarterly Journal of Economics*, 126(2):749–804, 2011.

Raj Chetty, John N Friedman, and Emmanuel Saez. Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review*, 103(7):2683–2721, 2013.

Federico Cingano and Alfonso Rosolia. People I know: Job Search and Social Networks. *Journal of Labor Economics*, 30(2):291–332, 2012.

Aaron Clauset, Mark EJ Newman, and Cristopher Moore. Finding Community Structure in Very Large Networks. *Physical review E*, 70(6):066111, 2004.

Guido Matias Cortes and Giovanni Gallipoli. The Costs of Occupational Mobility: An Aggregate Analysis. *Journal of the European Economic Association*, forthcoming.

Bruno Crepon, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment. *Quarterly Journal of Economics*, 128(2):531–580, 2013.

Leon Danon, Albert Diaz-Guilera, Jordi Duch, and Alex Arenas. Comparing Community Structure Identification. *Journal of Statistical Mechanics: Theory and Experiment*, 2005 (09):P09008, 2005.

Kunal Dasgupta. Learning and Knowledge Diffusion in a Global Economy. *Journal of International Economics*, 87(2):323–336, 2012.

Wolfgang Dauth, Sebastian Findeisen, and Jens Suedekum. The Rise of the East and the Far East: German Labor Markets and Trade Integration. *Journal of the European Economic Association*, 12(6):1643–1675, 2014.

Wolfgang Dauth, Sebastian Findeisen, and Jens Suedekum. Adjusting to Globalization - Evidence from Worker-Establishment Matches in Germany. Technical report, University of Mannheim, 2016.

Jed DeVaro. Employer Recruitment Strategies and the Labor Market Outcomes of New Hires. *Economic Inquiry*, 43(2):263–282, 2005.

Jed DeVaro. The Labor Market Effects of Employer Recruitment Choice. *European Economic Review*, 52(2):283–314, 2008.

Christian Dustmann and Costas Meghir. Wages, Experience and Seniority. *The Review of Economic Studies*, 72(1):77–108, 2005.

Christian Dustmann, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. Referral-based Job Search Networks. *The Review of Economic Studies*, 83(2):514–546, 2016a.

Christian Dustmann, Uta Schönberg, and Jan Stuhler. Labor Supply Shocks and the Dynamics of Local Wages and Employment. *Quarterly Journal of Economics*, forthcoming, 2016b.

Andreas Dzemski. An Empirical Model of Dyadic Link Formation in a Network with Unobserved Heterogeneity. Technical report, University of Mannheim, 2014.

Jonathan Eaton and Samuel Kortum. Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779, 2002.

B. Efron. Bootstrapping Methods: Another Look at the Jackknife. *Annals of Statistics*, 7: 1–26, 1979.

Glenn Ellison and Edward L. Glaeser. Geographic Concentration in US Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105:889–927, 1997.

Marc Ferracci, Grégory Jolivet, and Gerard J. van den Berg. Evidence of Treatment Spillovers Within Markets. *Review of Economics and Statistics*, 96(5):812–823, 2014.

Martina Fink, Esther Kalkbrenner, Andrea Weber, and Christine Zulehner. Extracting Firm Information from Administrative Records: The ASSD Firm Panel. Technical report, NRN Working Paper: The Austrian Center for Labor Economics and the Analysis of the Welfare State, 2010.

Santo Fortunato. Community Detection in Graphs. *Physics Reports*, 486(3):75–174, 2010.

Guido Friebel and Mariassunta Giannetti. Fighting for Talent: Risk-Taking, Corporate Volatility and Organisation Change. *The Economic Journal*, 119(540):1344–1373, 2009.

Xavier Gabaix and Augustin Landier. Why has CEO Pay Increased so Much? *The Quarterly Journal of Economics*, 123(1):49–100, 2008.

Manolis Galenianos. Learning about Match Quality and the Use of Referrals. *Review of Economic Dynamics*, 16(4):668–690, 2013.

Manolis Galenianos. Hiring through Referrals. *Journal of Economic Theory*, 152:304–323, 2014.

Christina Gathmann, Ines Helm, and Uta Schönberg. Spillover Effects in Local Labor Markets: Evidence from Mass Layoffs. Technical report, UCL, 2016.

Xavier Giroud and Holger M. Mueller. Firms□ Internal Networks and Local Economic Shocks. Technical report, MIT, 2017.

Albrecht Glitz. Coworker Networks in the Labour Market. *Labour Economics*, 44:218–230, 2017.

Bryan S. Graham. An Empirical Model of Link Formation with Degree Heterogeneity. *Econometrica*, forthcoming.

Michael Greenstone, Richard Hornbeck, and Enrico Moretti. Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, 118(3):536–598, 2010.

John Haltiwanger, Ron S Jarmin, and Javier Miranda. Who Creates Jobs? Small versus Large versus Young. *Review of Economics and Statistics*, 95(2):347–361, 2013.

John Haltiwanger, Henry Hyatt, and Erika McEntarfer. Cyclical Reallocation of Workers across Employers by firm size and firm wage. *National Bureau of Economic Research*, 2015.

Lena Hensvik and Oskar Nordström Skans. Social networks, employee selection and labor market outcomes. *Journal of Labor Economics*, 34(4):825–867, 2016.

Fredrik Heyman. Firm Size or Firm Age? The Effect on Wages using Matched Employer–employee Data. *Labour*, 21(2):237–263, 2007.

Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic Blockmodels: First Steps. *Social networks*, 5(2):109–137, 1983.

Arthur J Hosios. On the Efficiency of Matching and Related Models of Search and Unemployment. *The Review of Economic Studies*, 57(2):279–298, 1990.

Lawrence Hubert and Phipps Arabie. Comparing Partitions. *Journal of Classification*, 2(1): 193–218, 1985.

Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

Matthew O. Jackson and Brian W. Rogers. The Economics of Small Worlds. *Journal of the European Economic Association*, 3(2-3):617–627, 2005.

Hubert Janicki, Henry Hyatt, Emin Dinlersoz, et al. Who Works for Whom? Worker Sorting in a Model of Entrepreneurship with Heterogeneous Labor Markets. *Mimeo*, 2017.

Boyan Jovanovic. Job Matching and the Theory of Turnover. *Journal of Political Economy*, 87(5, Part 1):972–990, 1979.

Boyan Jovanovic. Matching, Turnover, and Unemployment. *Journal of Political Economy*, 92(1):108–122, 1984.

Brian Karrer and Mark E.J. Newman. Stochastic Blockmodels and Community Structure in Networks. *Physical Review E*, 83(1):016107, 2011.

Henrik J Kleven and Mazhar Waseem. Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics*, 128:669–723, 2013.

Henrik J Kleven, Claus T Kreiner, and Emmanuel Saez. Why can Modern Governments Tax so much? An Agency Model of Firms as Fiscal Intermediaries. *Economica*, 83(330):219–246, April 2016.

Henrik Jacobsen Kleven. Bunching. *Annual Review of Economics*, 8(1):435–464, 2016.

Francis Kramarz and Oskar Nordström Skans. When Strong Ties are Strong: Networks and Youth Labour Market Entry. *The Review of Economic Studies*, page rdt049, 2014.

Rafael Lalive. How do Extended Benefits Affect Unemployment Duration? A Regression Discontinuity Approach. *Journal of Econometrics*, 142(2):785 – 806, 2008.

Rafael Lalive and Josef Zweimüller. Benefit Entitlement and Unemployment Duration: The Role of Policy Endogeneity. *Journal of Public Economics*, 88(12):2587 – 2616, 2004.

Rafael Lalive, Camille Landais, and Josef Zweimüller. Market Externalities of Large Unemployment Extension Programs. *American Economic Review*, 105(12):3564–96, 2015.

Fabian Lange. The Speed of Employer Learning. *Journal of Labor Economics*, 25(1):1–35, 2007.

Edward Lazear and Paul Oyer. *Handbook of Organizational Economics*. Princeton University Press, December, 2012.

Edward P Lazear. Hiring Risky Workers. *National Bureau of Economic Research*, 1995.

Edward P Lazear and Mike Gibbs. *Personnel Economics in Practice*. John Wiley & Sons, 2014.

Michael Lechner, Conny Wunsch, and Patrycja Scioch. Do Firms Benefit from Active Labour Market Policies? Economics Working Paper Series 1318, University of St. Gallen, 2013.

James G. Mackinnon and Matthew D. Webb. Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics*, 2016. jae.2508.

Alan Manning and Barbara Petrongolo. How Local are Labor Markets? Evidence from a Spatial Job Search Model. *American Economic Review*, forthcoming.

D. M. Mason and M. A. Newton. A Rank Statistics Approach to the Consistency of a General Bootstrap. *The Annals of Statistics*, 20(3):1611–1624, 1992.

Enrico Moretti. Workers' Education, Spillovers, and Productivity: Evidence from Plant-level Production Functions. *American Economic Review*, 94(3):656–690, 2004.

Enrico Moretti. Local Labor Markets. In *Handbook of Labor Economics*, volume 4, pages 1237–1313. Elsevier, 2011.

Giuseppe Moscarini. Job Matching and the Wage Distribution. *Econometrica*, 73(2):481–516, 2005.

Giuseppe Moscarini and Fabien Postel-Vinay. The Timing of Labor Market Expansions: New Facts and a new Hypothesis. *NBER Macroeconomics Annual*, 23(1):1–52, 2008.

Giuseppe Moscarini and Fabien Postel-Vinay. Stochastic Search Equilibrium. *The Review of Economic Studies*, page rdt012, 2013.

Giuseppe Moscarini and Fabien Postel-Vinay. Did the Job Ladder Fail after the Great Recession? *Journal of Labor Economics*, 34(S1):S55–S93, 2016.

Éva Nagypál. Learning by Doing vs. Learning about Match Quality: Can we Tell them Apart? *The Review of Economic Studies*, 74(2):537–566, 2007.

Mark E.J. Newman and Michelle Girvan. Finding and Evaluating Community Structure in Networks. *Physical review E*, 69(2):026113, 2004.

Matthew J. Notowidigdo. The Incidence of Local Labor Demand Shocks. Technical report, National Bureau of Economic Research, 2013.

Paige Ouimet and Rebecca Zarutskie. Who Works for Startups? The Relation between Firm Age, Employee Age, and Growth. *Journal of Financial Economics*, 112(3):386–407, 2014.

Paul Oyer and Scott Schaefer. Personnel Economics: Hiring and Incentives. *National Bureau of Economic Research*, 2010.

Jörg Paetzold and Hannes Winner. Taking the High Road? Compliance with commuter tax allowances and the role of evasion spillovers. *Journal of Public Economics*, 143:1–14, 2016.

Tiago P Peixoto. Efficient Monte Carlo and Greedy Heuristic for the Inference of Stochastic Block Models. *Physical Review E*, 89(1):012804, 2014a.

Tiago P. Peixoto. Hierarchical Block Structures and High-resolution Model Selection in Large Networks. *Physical Review X*, 4(1):011047, 2014b.

Christopher A Pissarides. Short-run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages. *The American Economic Review*, 75(4):676–690, 1985.

Jennifer P. Poole. Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility. *Review of Economics and Statistics*, 95(2):393–406, 2013.

Michael Pries and Richard Rogerson. Hiring Policies, Labor Market Institutions, and Labor Market Flows. *Journal of Political Economy*, 113(4):811–839, 2005.

Emmanuel Saez. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, pages 180–212, 2010.

Perihan Ozge Saygin, Andrea Weber, and Michele A. Weynandt. Coworkers, Networks, and Job Search Outcomes. Technical report, IZA Discussion Paper 8174, 2014.

Ian M. Schmutte. Free to Move? A Network Analytic Approach for Learning the Limits to Job Mobility. *Labour Economics*, 29:49–61, 2014.

Michel Serafinelli. Good Firms, Worker Flows and Local Productivity. Technical report, University of Toronto, 2014.

Scott Shane. Why Encouraging more People to Become Entrepreneurs is Bad Public Policy. *Small Business Economics*, 33(2):141–149, 2009.

Isaak Sorkin. Ranking Firms Using Revealed Preference. Technical report, Unpublished Manuscript, 2015.

Andrey Stoyanov and Nikolay Zubanov. Productivity Spillovers across Firms through Worker Mobility. *American Economic Journal: Applied Economics*, 4(2):168–198, 2012.

Andrey Stoyanov and Nikolay Zubanov. The Distribution of the Gains from Spillovers through Worker Mobility between Workers and Firms. *European Economic Review*, 70: 17–35, 2014.

Marko Terviö. The Difference that CEOs Make: An Assignment Model Approach. *The American Economic Review*, 98(3):642–668, 2008.

Yuchung J. Wang and George Y. Wong. Stochastic Blockmodels for Directed Graphs. *Journal of the American Statistical Association*, 82(397):8–19, 1987.

Xiaoran Yan, Cosma Shalizi, Jacob E. Jensen, Florent Krzakala, Cristopher Moore, Lenka Zdeborova, Pan Zhang, and Yaojia Zhu. Model Selection for Degree-corrected Block Models. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(5):P05007, 2014.

Yunpeng Zhao, Elizaveta Levina, and Ji Zhu. Consistency of Community Detection in Networks under Degree-corrected Stochastic Block Models. *The Annals of Statistics*, 40(4): 2266–2292, 2012.

Josef Zweimüller, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philippe Wuellrich, Oliver Ruf, and Simon Büchi. Austrian Social Security Database. *NRN: The Austrian Center for Labor Economics and the Analysis of the Welfare State*, 2009.